

LIBERA UNIVERSITÀ INTERNAZIONALE DEGLI STUDI SOCIALI

“LUISS - GUIDO CARLI”

LUISS



DIPARTIMENTO DI ECONOMIA E FINANZA

Indirizzo Banche e Intermediari Finanziari

**RELAZIONE TRA IL VOLUME DI
TRADING E LA CORRELAZIONE SERIALE
DEI RENDIMENTI AZIONARI**

RELATORE:

PROF. BORRI NICOLA

CORRELATORE:

PROF. BARONE EMILIO

LAUREANDO:

MEMA ELDI

ANNO ACCADEMICO 2023-2024

*Il vento spegne la candela e alimenta il fuoco.
Lo stesso vale per il caso, l'incertezza e il caos:
vogliamo usarli, non nasconderci da loro.
Vogliamo essere fuoco e desiderare il vento.*

Indice

Indice

Elenco delle tabelle

Elenco delle figure

Introduzione	2
1 Analisi delle Variabili	4
1.1 Analisi dei rendimenti	6
1.1.1 Definizione dei Log-Rendimenti	6
1.1.2 Vantaggi dei Log-Rendimenti	6
1.1.3 Correlazione seriale	7
1.1.4 Risultati e Interpretazioni	7
1.1.5 Illustrazioni grafiche	8
1.2 Analisi del Volume di Trading	8
1.2.1 Vantaggi dell'Utilizzo del Log Volume Detrended	10
1.2.2 Logaritmo	11
1.2.3 Media Mobile Triangolare	11
1.2.4 Volume Detrended	11
1.2.5 Stazionarietà	12
1.3 Pulizia dei Dati con il Filtro di Hampel	12
1.3.1 Descrizione del Filtro di Hampel	12
1.3.2 Perché non altre metriche?	13
1.4 Analisi della volatilità: Il Modello GARCH	14
1.4.1 Volatilità	14
1.4.2 Volatilità Condizionata	15
1.4.3 Modello GARCH	15
2 Previsione dei rendimenti	18
2.0.1 Variabili Dummy	19
2.1 Autocorrelazione dei rendimenti dell'S&P 500	20
2.1.1 Autocorrelazione dei Rendimenti in Relazione a Volume, Volatilità e Dummies	21
2.1.2 Seconda autocorrelazione e la relazione con le variabili	24
2.1.3 Autocorrelazione di secondo ordine	24
2.1.4 Autocorrelazione di secondo ordine e la relazione con il volume e la volatilità	25
2.2 Modello alternativo: la MA dei rendimenti e Volume	27

2.2.1	Modello con Media Mobile dei Rendimenti Passati	27
2.2.2	Modello con Media Mobile Interagita con il Volume	27
2.2.3	Commenti sui Risultati delle Regressioni	28
2.3	Modello a Componenti Non Osservate	29
2.3.1	Specificazione del Modello Locale	29
2.3.2	Modello a Componenti non Osservate: Complesso	30
2.4	L'Utilizzo del Rendimento Quadratico nella Regressione	32
2.5	Modello con V_t detrendizzato a 3m, 6m e 3y	33
2.5.1	Risultati con Medie Mobili Triangolari differenti	33
2.6	Sviluppi Intertemporali: La Ricerca di Campbell	36
2.7	Modello di Markov Switching	40
2.7.1	Descrizione del Modello	41
2.7.2	Equazioni del Modello	41
2.7.3	Risultati del Modello	41
2.7.4	Regimi Identificati	42
2.7.5	Probabilità di Transizione	42
2.7.6	Impatto del Volume e Confronto con i Risultati Precedenti	43
2.8	Trading Non Sincrono	44
2.8.1	Impatto del Trading Non Sincrono sulla Correlazione Seriale dei Rendimenti	44
2.8.2	Modello alternativo per il Trading non Sincrono: Analisi dei rendimenti individuali	48
3	Volume e Correlazione: Un Modello Teorico	52
3.1	Struttura dell'Economia	52
3.1.1	Processo Stocastico dei Dividendi	53
3.1.2	Tipi di Investitori	53
3.2	Utilità Attesa, Prezzo e Segnale	54
3.2.1	Massimizzazione dell'Utilità Attesa	54
3.2.2	Set Informativo e Segnale	55
3.2.3	Prezzo di Equilibrio dell'Asset Rischioso	55
3.2.4	Varianza dell'Innovazione del Prezzo	55
3.3	Prezzo e Avversione al rischio	57
3.3.1	Definizione della Variabile Z_t	57
3.3.2	Processo AR(1) per \tilde{Z}_t	58
3.3.3	Assunzioni sul Processo Z_t	58
3.3.4	Prezzo di Equilibrio dell'Asset	58
3.4	Rendimento in Eccesso e Volume	59
3.4.1	Rendimento in Eccesso	59
3.4.2	Frazioni Ottimali di Azioni Detenute da Investitori di Tipo A e B	60
3.4.3	Volume di Trading V_t	62
3.4.4	Implicazioni del Modello per la Correlazione tra Volume e Rendimento Seriale	63
3.4.5	Predizione dei Ritorni Futuri con Z_t	63
3.4.6	Modello non lineare con le variabili del Volume di Trading e il rendimento	64
3.4.7	Relazione tra Volume e Rendimento	65
3.5	Simulazione Del Modello	66

3.5.1	Calibrazione del Coefficiente λ e della Varianza σ_Z^2	68
3.5.2	Definizione dei parametri e Analisi dei Risultati	68
3.5.3	Implicazioni Empiriche del Modello	71
	Conclusione	73
	Appendice	74
	Codici Python	80
	Bibliografia	116

Elenco delle tabelle

1.1	Autocorrelazione e Statistiche dei Log-Rendimenti per Diversi Periodi . . .	7
1.2	Autocorrelazione e Statistiche del Log Volume Detrendizzato per Diversi Periodi	14
2.1	La Prima Autocorrelazione dei Rendimenti Azionari	20
2.2	Volume, Volatilità, e la Prima Autocorrelazione dei Rendimenti	22
2.3	Seconda Autocorrelazione dei Rendimenti Azionari	24
2.4	Volume, Volatilità, e la Seconda Autocorrelazione dei Rendimenti	26
2.5	Media Mobile e Volume nella Previsione dei Rendimenti	28
2.6	Volume e la Prima Autocorrelazione dei Rendimenti: Modello a Componenti Locali e Non Osservate	32
2.7	Volume, Volatilità, e la Prima Autocorrelazione dei Rendimenti con Rendimento Quadratico	34
2.8	Volume e la Prima Autocorrelazione dei Rendimenti: Modello con Medie Mobili Triangolari	35
2.9	La Prima Autocorrelazione Dei Rendimenti Azionari: Un Campionamento Alternativo	37
2.10	Volume, Volatilità, e la Prima Autocorrelazione dei Rendimenti: Un Campionamento Alternativo	38
2.11	Risultati del Modello di Markov Switching	42
2.12	Parametri di Transizione del Modello di Markov Switching	42
2.13	Beta e R-Squared: Autocorrelazione di Primo Ordine del Dow Jones Industrial Average	45
2.14	Volume, Volatilità, e la Prima Autocorrelazione dei Rendimenti: Dow Jones Industrial Average	47
2.15	Effetto del Volume e delle Variabili Temporal sui Rendimenti Azionari	50
3.1	Simulazioni di Volume e Autocorrelazione di Primo Ordine	69
3.2	Statistiche di Rendimento, Volume e Z_t per diversi valori di α_Z	71

Elenco delle figure

1.1	Andamento del prezzo dell'S&P 500 dal 1988 al 2023	8
1.2	Autocorrelazione dei rendimenti dell'S&P 500 dal 1988 al 2023	9
1.3	Andamento del Volume dell'S&P 500 dal 1988 al 2023	10
1.4	Evoluzione del log Volume detrendizzato dell'S&P 500 dal 1988 al 2023 .	13
1.5	Confronto tra volatilità condizionata e deviazione standard mobile del S&P 500	16
2.1	Panoramica sui regimi ed i rendimenti dello S&P 500	43
3.1	Simulazioni di Z_t e b_t con i vari valori di α_Z	70

Introduzione

L'analisi delle dinamiche del mercato azionario ha evidenziato come il volume degli scambi e la volatilità dei rendimenti possano fornire preziose informazioni sulla struttura e sull'efficienza dei mercati finanziari. Tuttavia, la relazione tra il volume di trading e la correlazione seriale dei rendimenti azionari è stata oggetto di attenzione limitata in molti studi economici. Questo lavoro si propone di aggiornare e rielaborare alcune delle intuizioni chiave proposte nel lavoro seminale di *Campbell, Grossman e Wang (1993)*¹, in cui gli autori esploravano l'influenza del volume degli scambi sulla correlazione seriale dei rendimenti azionari, con particolare attenzione ai meccanismi di mercato legati agli investitori non informativi e alla risposta dei market maker avversi al rischio. Il lavoro di *Campbell et al.* rimane una pietra miliare nella letteratura economica, evidenziando come le giornate caratterizzate da un alto volume di scambi presentino una bassa autocorrelazione dei rendimenti rispetto a quelle a basso volume. Tuttavia, i mercati finanziari hanno subito numerosi cambiamenti nel tempo, tra cui l'introduzione di nuove tecnologie di trading e l'aumento dell'influenza degli investitori istituzionali. Eventi di rilevanza globale, come la crisi finanziaria del 2008 e la pandemia da COVID-19, hanno ulteriormente modificato le dinamiche di mercato. In questo contesto, rielaborare e aggiornare il modello proposto da *Campbell et al.* diventa cruciale per comprendere meglio le caratteristiche del mercato odierno.

L'intuizione alla base di questo lavoro è la seguente: quando si osserva una diminuzione dei prezzi azionari, essa può essere attribuita a due cause principali. Nel primo caso, una nuova informazione pubblica potrebbe portare tutti gli investitori a rivedere al ribasso le proprie valutazioni sul mercato azionario. In alternativa, tale riduzione potrebbe derivare da una pressione di vendita esogena esercitata da investitori non informativi. Nel caso di informazioni pubbliche, non ci si aspetta una variazione significativa nel rendimento atteso del mercato azionario. Tuttavia, quando la pressione di vendita proviene da operatori non informativi, i market maker richiedono un rendimento atteso più elevato per accomodare tale vendita, portando probabilmente a un aumento dei prezzi nei giorni successivi. Questi due scenari possono essere distinti osservando il volume di scambi: se le fluttuazioni sono guidate da informazioni pubbliche, il volume non dovrebbe aumentare in modo significativo; viceversa, un volume elevato di scambi potrebbe indicare una pressione di vendita da parte di investitori non informativi.

Il modello con investitori eterogenei suggerisce inoltre che i cambiamenti di prezzo accompagnati da volumi elevati tendono a invertirsi nei giorni successivi, mentre le variazioni di prezzo in giornate a basso volume non mostrano la stessa tendenza. Le variazioni nella domanda di azioni da parte degli operatori non informativi possono verificarsi sia a bassa che ad alta frequenza. Il volume degli scambi giornalieri è un indicatore di variazioni ad alta frequenza della domanda. Tuttavia, cambiamenti che si verificano lentamente nel

¹Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). "Trading Volume and Serial Correlation in Stock Returns." *The Quarterly Journal of Economics*, 108(4), 905-939.

tempo sono più difficili da rilevare attraverso i dati sul volume, poiché il volume può essere influenzato da altre dinamiche di mercato, come la deregolamentazione delle commissioni e la crescita degli scambi istituzionali.

Questo lavoro si distingue dai precedenti aggiornando tali intuizioni alla luce di eventi recenti e sviluppi tecnologici nei mercati finanziari. Un aspetto particolarmente rilevante che questo studio si propone di approfondire è il comportamento degli investitori “liquidi” o non informativi e la loro interazione con i market maker avversi al rischio, in particolare nei periodi di crisi economica, come la crisi sub-prime e il crollo del 2020 dovuto al COVID-19. Durante questi periodi, si è osservata un’evidenza empirica di una maggiore relazione tra il volume degli scambi e i rendimenti azionari, con un incremento dell' R^2 nelle regressioni e una maggiore incidenza del coefficiente legato al volume di scambi. Questo suggerisce che la pressione esercitata dagli investitori non informativi si amplifica nelle fasi di crisi, con effetti osservabili sull’autocorrelazione dei rendimenti e sulla volatilità di mercato.

L’organizzazione di questo lavoro è definito nel seguente modo: nel **Capitolo 1**, verranno introdotte e analizzate le principali variabili di studio, ossia rendimento, volume di trading e volatilità, fornendo un quadro teorico e descrivendo le loro interazioni. Nel **Capitolo 2**, l’attenzione si sposterà sulla previsione dei rendimenti, attraverso l’analisi delle relazioni tra le variabili tramite diverse metodologie, partendo da semplici relazioni lineari fino ad arrivare a modelli di regressione più complessi, incluse le componenti non osservate, e utilizzando un campione dei principali titoli azionari nel periodo dal 2004 al 2023. Infine, nel **Capitolo 3**, verrà presentato il modello teorico utilizzato per simulare il comportamento di mercato, spiegandone le caratteristiche e giungendo alla realizzazione di un modello simulativo basato su tali teorie.

Capitolo 1

Analisi dei Rendimenti, Volume di trading e Volatilità

In questo elaborato, i principali dati finanziari utilizzati sono quelli dello Standard & Poor's 500 e, a seguire, quelli del Dow Jones e dei principali 30 titoli azionari statunitensi del periodo di campionamento. Per quanto concerne lo Standard & Poor's 500, è uno degli indici borsistici più seguiti al mondo e rappresenta un indicatore chiave della performance delle azioni statunitensi. L'indice comprende 500 delle maggiori società quotate sulle borse americane, principalmente il New York Stock Exchange (NYSE) e il NASDAQ¹. Le aziende incluse nell'S&P 500 sono selezionate in base alla loro capitalizzazione di mercato, liquidità e settore industriale, rendendo l'indice una rappresentazione ampia e diversificata del mercato azionario statunitense. L'S&P 500 è comunemente utilizzato come benchmark per confrontare la performance di portafogli d'investimenti e fondi comuni, nonché come base per prodotti finanziari come fondi indicizzati e contratti futures². La diversificazione dell'indice, che copre tutti i principali settori economici, lo rende uno strumento affidabile per analizzare le tendenze generali del mercato azionario e dell'economia statunitense³.

Per questo motivo, data la necessità di comprendere la relazione tra l'impatto del volume sull'autocorrelazione dei rendimenti, la scelta di utilizzare questo indice come principale serie di rendimenti giornalieri è la più appropriata. In questo senso, i vari dati sul volume di scambi su cui si sono svolte le principali ricerche e analisi dell'indice, si intendono come la somma dei volumi delle varie società che compongono l'indice e che sono state scambiate nel NYSE e nel NASDAQ.

L'intervallo temporale su cui si sono svolte principalmente le analisi parte dal 1 gennaio 2004 al 31 dicembre 2023 per diverse ragioni significative:

1. **Copertura di eventi economici rilevanti:** Include la crisi finanziaria globale del 2008-2009 e la pandemia di COVID-19 del 2020-2021, permettendo di studiare l'impatto di shock significativi sui mercati finanziari.

¹Wikipedia contributors. (2023). S&P 500. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/S%26P_500

²Forbes, B. (2024). Come investire nello S&P 500. Retrieved from <https://www.forbes.com/advisor/it/investire/sp500-come-investire-in-borsa/>

³Yahoo Finance contributors. (2023). What is the S&P 500 and How Does it Work? In Yahoo Finance. Retrieved from <https://finance.yahoo.com/news/what-is-the-s-p-500-123456789.html>

2. **Ampia finestra temporale:** Un periodo di vent'anni consente di osservare tendenze a lungo termine e ciclicità del mercato, fornendo un'analisi più completa e rappresentativa.
3. **Evoluzione delle politiche regolamentari:** Copre importanti cambiamenti regolamentari post-crisi, che hanno influenzato i comportamenti degli investitori e la struttura del mercato soprattutto a partire dalla crisi sub-prime in cui si sono rivisitate le varie regolamentazioni di trasparenza e correttezza.

Data la presenza delle due grandi crisi presenti nella ricerca, potrebbe portare a un dominio di quest'ultime sui dati giornalieri che possiamo osservare. Per questo motivo, il periodo di campionamento è stato suddiviso in sottocampioni, per comprendere le varie dinamiche presenti nel periodo interessato. I periodi di sottocampionamento sono così definiti:

- **Full Period:** 01/01/2004 - 31/12/2023
- **Crisi Sub-prime:** 01/01/2005 - 31/12/2009
- **Post crisi subprime:** 01/01/2010 - 31/12/2012
- **Ripresa economica:** 01/01/2013 - 31/12/2019
- **Covid-19 e Guerre:** 01/01/2020 - 31/12/2023

Oltre allo studio in merito ai rendimenti dello S&P 500, è stato studiato anche il comportamento di altre serie di dati finanziari. Una prima analisi è stata fatta sempre in merito allo S&P 500 ma con un campionamento temporale definito dal 01/01/1988 al 31/12/2023. Tale scelta è spiegata dalle evoluzioni regolamentari presenti nell'intervallo e l'efficienza che ne è derivata sui mercati finanziari.

Altresì è stato studiato il comportamento del Dow Jones per il periodo compreso dal 2004 al 2023 e, in linea con l'analisi svolta per lo S&P 500, dal periodo 1988 al 2023. Il perché dietro questo secondo studio è dato dalle dimensioni delle società che compongono l'indice. In particolare, quest'indice è composto dalle 30 maggiori e più influenti società industriali quotate nelle borse statunitensi, scelte dal Wall Street Journal. Queste aziende coprono una vasta gamma di settori economici, ad eccezione dei trasporti e delle utility, che sono rappresentati da altri indici Dow Jones specifici. A differenza dello S&P 500, questo indice è ponderato per il prezzo. Ciò significa che le società con prezzi delle azioni più elevati hanno un maggiore impatto sul valore dell'indice ⁴.

L'utilizzo di questo indice è di particolare importanza perché il comportamento dei titoli di grandi dimensioni è di particolare interesse, poiché è improbabile che i rendimenti misurati di questi titoli siano influenzati da un trading non sincrono. Anche i rendimenti dei singoli titoli forniscono un'utile evidenza robusta al trading non sincrono, pertanto viene studiato anche il comportamento dei rendimenti dei 30 grandi titoli del periodo compreso dal 2004 al 2023.

In linea con la ricerca svolta da Campbell, Grossman e Wang, questa tesi vuole rendersi partecipe dell'evidenza empirica dei risultati ottenuti da quest'ultimi. Per far ciò, sono stati utilizzati dati principalmente provenienti da Bloomberg e dalla piattaforma di Yahoo Finance.

⁴Investopedia. (2023). What Does the Dow Jones Index Measure and How Is It Calculated? Retrieved from <https://www.investopedia.com/terms/s/DJ.asp>

1.1 Analisi dei rendimenti

Per questa ricerca sono stati utilizzati i log-rendimenti per comprendere il loro andamento nel tempo e successivamente sono state calcolate le autocorrelazioni a 1 e 5 lag per comprendere la reazione dei rendimenti rispetto a quelli passati. Infine, si presentano i risultati di questa analisi attraverso una tabella esauriente di risultati e grafici che rappresentano l'andamento del prezzo del rendimento e delle autocorrelazioni.

1.1.1 Definizione dei Log-Rendimenti

Prima di spiegare cosa sia e perché sia stato utilizzato il log-rendimento, è necessario definire cosa sia alla base il rendimento. Il rendimento di un'attività finanziaria è una misura della variazione relativa del suo prezzo rispetto a un periodo di riferimento, ed è utilizzato per quantificare la performance dell'investimento. Esso rappresenta la percentuale di guadagno o perdita rispetto al valore iniziale e può essere calcolato come la differenza tra il prezzo finale e il prezzo iniziale, rapportata al prezzo iniziale stesso⁵.

I log-rendimenti, invece, sono una trasformazione logaritmica dei rendimenti semplici. Se P_t rappresenta il prezzo di un'attività finanziaria al tempo t , il rendimento semplice R_t tra il tempo t e $t - 1$ è dato da:

$$R_t = \frac{P_t}{P_{t-1}} - 1$$

Il log-rendimento r_t è invece definito come:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln(P_t) - \ln(P_{t-1})$$

1.1.2 Vantaggi dei Log-Rendimenti

L'utilizzo dei log-rendimenti presenta diversi vantaggi rispetto ai rendimenti semplici:

1. **Additività Temporale:** I log-rendimenti sono additivi nel tempo. Se si considerano più periodi consecutivi, il log-rendimento totale è semplicemente la somma dei log-rendimenti di ciascun periodo:

$$r_{\text{totale}} = \sum_{t=1}^n r_t$$

2. **Simmetria:** I log-rendimenti trattano aumenti e diminuzioni dei prezzi in modo simmetrico. Un aumento del 10% seguito da una diminuzione del 10% non restituisce il prezzo iniziale con i rendimenti semplici, mentre con i log-rendimenti la somma algebrica torna a zero.

3. **Normalità:** I log-rendimenti tendono ad essere più prossimi ad una distribuzione normale rispetto ai rendimenti semplici, facilitando l'applicazione di strumenti statistici tradizionali che assumono la normalità dei dati⁶.

⁵Damodaran, A. (2020). *Investment Valuation: Tools and Techniques for Determining the Value of Any Asset* (3rd ed.). Wiley.

⁶DSPyt. (2022). How to illustrate log returns vs simple returns. Retrieved from <https://dspyt.com/simple-returns-log-return-and-volatility-simple-introduction>

4. **Linearità:** I log-rendimenti rendono lineari i dati che tendono ad avere un andamento esponenziale. Questo significa che, per processi di crescita esponenziale, i log-rendimenti possono semplificare l'analisi e la modellizzazione, in quanto trasformano una curva esponenziale in una linea retta⁷.

1.1.3 Correlazione seriale

La correlazione seriale o autocorrelazione, misura il grado di correlazione di una serie temporale con sé stessa a diversi ritardi. Se r_t rappresenta il log-rendimento al tempo t , l'autocorrelazione al ritardo k è definita come:

$$\rho_k = \frac{\text{Cov}(r_t, r_{t+k})}{\sigma_{r_t} \sigma_{r_{t+k}}}$$

dove:

- $\text{Cov}(r_t, r_{t+k})$ è la covarianza tra r_t e r_{t+k} ,
- σ_{r_t} e $\sigma_{r_{t+k}}$ sono le deviazioni standard di r_t e r_{t+k} .

L'autocorrelazione può essere stimata utilizzando la formula:

$$\hat{\rho}_k = \frac{\sum_{t=1}^{T-k} (r_t - \bar{r})(r_{t+k} - \bar{r})}{\sum_{t=1}^T (r_t - \bar{r})^2}$$

dove:

- T è il numero di osservazioni,
- \bar{r} è la media dei log-rendimenti.

1.1.4 Risultati e Interpretazioni

I risultati delle autocorrelazioni dei rendimenti logaritmicici per i diversi periodi di sottocampionamento sono riassunti nella tabella seguente:

Tabella 1.1: Autocorrelazione e Statistiche dei Log-Rendimenti per Diversi Periodi

Periodo	ρ_1	ρ_5	\bar{x}	σ
Crisi Sub-Prime	-0.133	-0.025	0.000	0.015
Post Crisi Subprime	-0.072	-0.126	0.000	0.012
Ripresa Economica	-0.027	-0.034	0.000	0.008
Covid-19 e Guerre	-0.196	0.090	0.000	0.015
Full Period	-0.123	-0.006	0.000	0.012

Interpretazione dei Risultati

⁷DSPyt. (2022). How to illustrate log returns vs simple returns. Retrieved from <https://dspyt.com/simple-returns-log-return-and-volatility-simple-introduction>

La tabella mostra autocorrelazioni negative nei rendimenti a breve termine (ρ_1) per la maggior parte dei periodi, indicando una tendenza all'inversione dei rendimenti. Tuttavia, a lungo termine (ρ_5), le correlazioni sono più vicine allo zero o leggermente positive, suggerendo una stabilizzazione dei rendimenti nei periodi successivi. La volatilità (σ) rimane relativamente stabile con valori bassi in tutti i periodi, tranne nei momenti di maggiore incertezza economica.

Questi risultati evidenziano come i fattori psicologici e le reazioni emotive influenzano le decisioni di investimento, portando a comportamenti di overreaction durante i periodi di crisi e stabilizzazione durante i periodi di ripresa economica.

Per comprendere la realtà di cui trattiamo si evidenzia attraverso i grafici l'andamento del prezzo ed i relativi periodi di crisi e l'andamento dell'autocorrelazione.

1.1.5 Illustrazioni grafiche

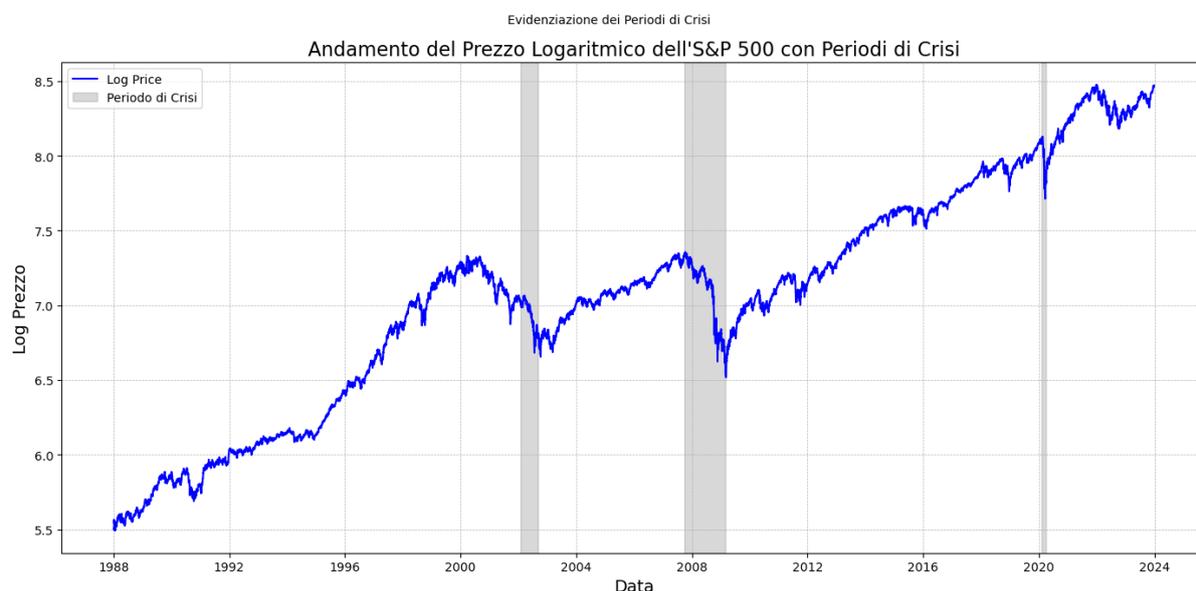


Figura 1.1: Andamento del prezzo dell'S&P 500 dal 1988 al 2023

Come si evince dal grafico, l'indice S&P 500 ha avuto un andamento, visto in scala logaritmica, crescente con qualche periodo di crisi di mercato.

Per quanto concerne l'autocorrelazione dei rendimenti, si illustra il suo andamento con lag da 1 a 250 (giorni di borsa) nella figura 2.

1.2 Definizione e Analisi del Volume di Trading

Il volume di trading rappresenta il numero totale di azioni, contratti o lotti scambiati su un determinato strumento finanziario in un periodo di tempo specifico, come un giorno, una settimana o un mese in un determinato mercato. Questo parametro fornisce informazioni cruciali sulla liquidità di un'attività e sull'interesse degli investitori. Un volume elevato può indicare un forte interesse per l'attività e spesso accompagna movimenti significativi

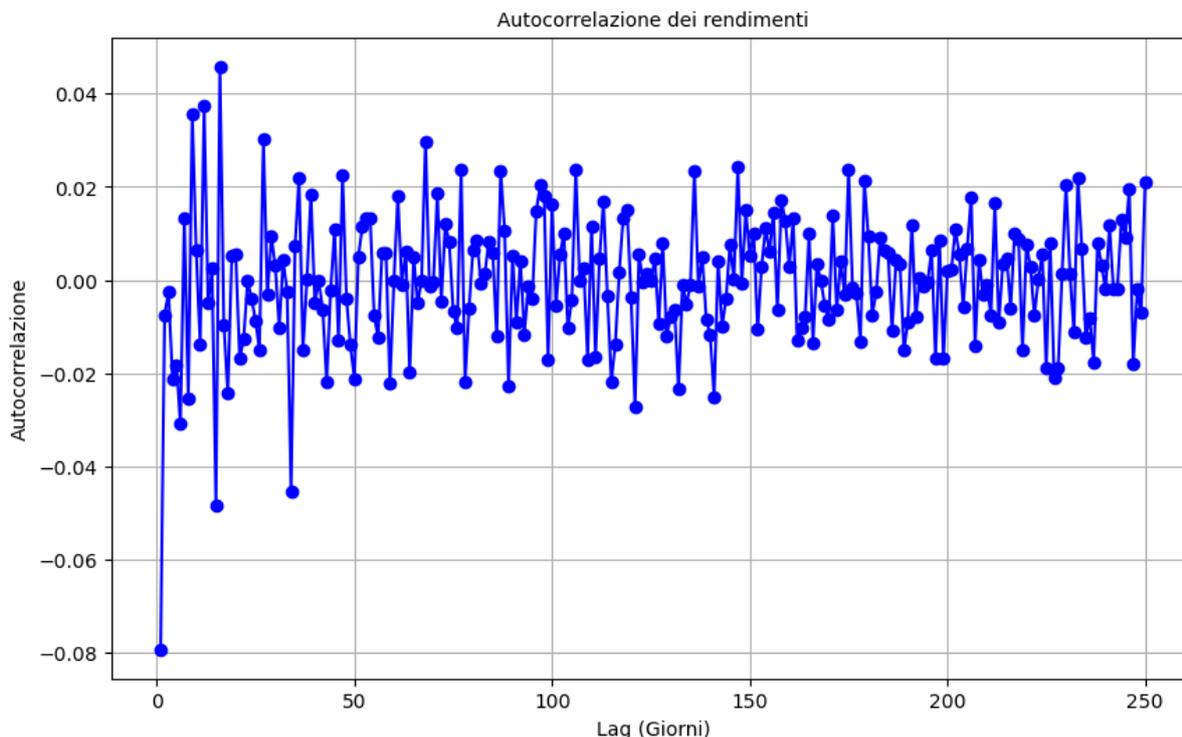


Figura 1.2: Autocorrelazione dei rendimenti dell'S&P 500 dal 1988 al 2023

nei prezzi, sia al rialzo che al ribasso. Inoltre, il volume è spesso utilizzato come indicatore complementare nelle analisi tecniche per confermare tendenze di mercato⁸.

Il volume di trading non è solo una misura della quantità di scambi, ma anche un indicatore di consenso o dissenso tra gli investitori riguardo al valore dell'attività. Un aumento improvviso del volume può suggerire un cambiamento nelle aspettative del mercato o l'ingresso di nuovi operatori nel mercato⁹. Questo lo rende un elemento fondamentale per comprendere la dinamica dei mercati e il comportamento degli investitori.

Per questo elaborato il volume ha un peso cruciale per le varie analisi che ne seguono, per cercare di avere quanto più dati veritieri e corretti possibili, il volume di scambi e i prezzi sono stati scaricati dalla piattaforma Bloomberg. Tuttavia, i dati così presentati sul volume presentano alcune criticità che possono compromettere l'analisi statistica e l'interpretazione dei risultati:

1. **Trend di lungo periodo:** Il volume degli scambi tende a mostrare trend di lungo periodo dovuti a fattori esterni come la crescita economica, l'aumento del numero di investitori e l'espansione delle attività di trading. Questi trend possono introdurre una componente spuria nei dati, rendendo difficile isolare le variazioni rilevanti del volume.
2. **Estremi e Outliers:** I dati sul volume degli scambi possono contenere valori estremi o outliers che distorcono la distribuzione e possono influenzare negativamente le analisi statistiche. Gli outliers possono derivare da eventi eccezionali o irregolari.

⁸Harris, L. (2003). *Trading and Exchanges: Market Microstructure for Practitioners*. Oxford University Press.

⁹Karpoff, J. M. (1987). The Relation Between Price Changes and Trading Volume: A Survey. *The Journal of Financial and Quantitative Analysis*, 22(1), 109-126.

3. **Stagionalità e ciclicità:** Il volume degli scambi può essere influenzato da effetti stagionali o ciclici che non sono necessariamente correlati alle variazioni di mercato di interesse.

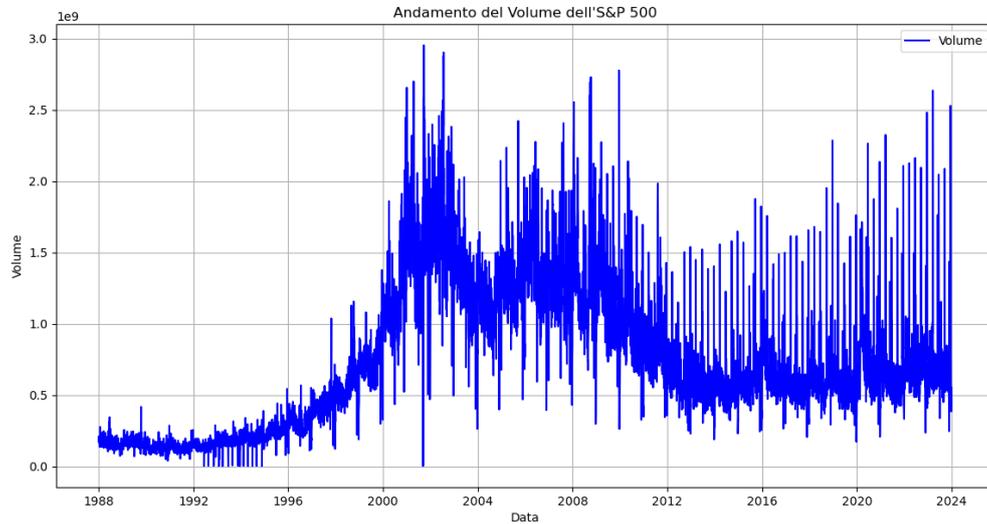


Figura 1.3: Andamento del Volume dell'S&P 500 dal 1988 al 2023

Per tali ragioni, è stato deciso di utilizzare il volume in logaritmo detrendizzato per una media mobile triangolare con un intervallo temporale di 250 giorni. Successivamente è stato svolto uno studio con un intervallo di 3 mesi, 6 mesi e 3 anni per comprendere la robustezza delle analisi, è stato svolto altresì uno studio con il volume a componenti non osservate.

1.2.1 Vantaggi dell'Utilizzo del Log Volume Detrended

Detrendizzazione

La detrendizzazione è il processo di rimozione dei trend a lungo termine dai dati per evidenziare le variazioni cicliche o stagionali¹⁰. In questa ricerca è stata utilizzata la media mobile triangolare (TMA) per detrendizzare il volume degli scambi. La formula della media mobile triangolare per una serie temporale V_t con una finestra di lunghezza $2N + 1$ è:

$$\text{TMA}(V_t) = \frac{1}{N(N+1)} \sum_{i=-N}^N (N+1-|i|)V_{t+i}$$

dove:

- V_t è il volume degli scambi al tempo t .
- N è il numero di periodi che determina la metà della finestra di osservazione.

¹⁰Statistics How To. (2023). Detrend Data. Retrieved from <https://www.statisticshowto.com/detrend-data/>

- i è l'indice che varia da $-N$ a N .
- $(N + 1 - |i|)$ è il peso assegnato a ciascun termine nella finestra.

11

1.2.2 Logaritmo

L'applicazione del logaritmo al volume degli scambi ha diversi benefici:

- **Normalizzazione:** Il logaritmo trasforma la distribuzione del volume, riducendo l'asimmetria e avvicinandola a una distribuzione normale.
- **Riduzione della Variabilità:** Il logaritmo attenua l'impatto degli outliers, rendendo la serie temporale più robusta e meno sensibile a valori estremi.
- **Proprietà Additive:** Le proprietà matematiche del logaritmo rendono più facile interpretare le relazioni proporzionali e moltiplicative.
- **Linearizzazione:** Il logaritmo consente di trasformare una relazione esponenziale in una lineare. Ad esempio, se il volume segue una crescita esponenziale del tipo $V_t = e^{kt}$, allora $\log(V_t) = kt$, rendendo possibile l'utilizzo di tecniche di regressione lineare.

La trasformazione logaritmica di una serie temporale del volume V_t è data da:

$$\log(V_t)$$

1.2.3 Media Mobile Triangolare

L'utilizzo di una media mobile triangolare offre una pesatura simmetrica che dà più importanza ai dati centrali e meno ai dati più lontani, rendendo la serie temporale più liscia e riducendo la varianza indotta dagli outliers¹².

1.2.4 Volume Detrended

Il volume detrended è ottenuto sottraendo la media mobile triangolare dalla serie logaritmica del volume. Formalmente:

$$\text{LogVolumeDetrended}_t = \log(V_t) - \text{TMA}(\log(V_t))$$

dove:

- $\text{LogVolumeDetrended}_t$ è il volume logaritmico detrendizzato al tempo t .
- $\log(V_t)$ è il volume logaritmico al tempo t .
- $\text{TMA}(\log(V_t))$ è la media mobile triangolare del volume logaritmico¹³.

¹¹Arxiv.org. (2023). Effects of polynomial trends on detrending moving average analysis. Retrieved from <https://arxiv.org/abs/1505.02750>

¹²Duttaviews.com. (2023). Understanding the Triangular Moving Average (TMA) Indicator for Technical Analysis. Retrieved from <https://duttaviews.com/triangular-moving-average/>

¹³Duttaviews.com. (2023). Understanding the Triangular Moving Average (TMA) Indicator for Technical Analysis. Retrieved from <https://duttaviews.com/triangular-moving-average/>

1.2.5 Stazionarietà

Rendere una serie temporale stazionaria è fondamentale per molte tecniche di analisi dei dati e modelli econometrici. Una serie stazionaria ha proprietà statistiche, come la media e la varianza, che non cambiano nel tempo, rendendo i risultati delle analisi più affidabili e interpretabili¹⁴. La detrendizzazione e la trasformazione logaritmica contribuiscono entrambe a rendere la serie stazionaria.

1.3 Pulizia dei Dati con il Filtro di Hampel

Nel processo di analisi dei dati finanziari, è emersa la necessità di pulire i dati dei volumi delle aziende e degli indici per ottenere una serie temporale quanto più stazionaria possibile. Per questa pulizia è stato utilizzato il filtro di Hampel per individuare e correggere gli outlier, i quali possono derivare da eventi straordinari o anomalie nei dati.

1.3.1 Descrizione del Filtro di Hampel

Il filtro di Hampel è una tecnica robusta per la rilevazione degli outlier, che si basa sull'utilizzo della mediana anziché della media per individuare valori anomali. Come descritto da Li e Hu (2017) nel loro studio sull'applicazione dei filtri robusti in finanza, "il filtro di Hampel è particolarmente efficace nell'identificare valori anomali in dati finanziari rumorosi, migliorando così l'accuratezza delle previsioni econometriche"¹⁵. La robustezza del filtro deriva dal fatto che la mediana è meno influenzata dagli outlier rispetto alla media. In particolare, il filtro di Hampel funziona nel seguente modo:

- **Finestra Mobile:** Viene utilizzata una finestra mobile di dimensione w per scorrere lungo la serie temporale, in questo documento viene utilizzata una finestra di 66 giorni.
- **Mediana:** Per ogni punto x_t nella finestra, si calcola la mediana M_t dei valori contenuti nella finestra.
- **Deviazione Standard:** Successivamente, si calcola la deviazione standard σ_t all'interno della finestra.
- **Soglia di Outlier:** Un punto x_t viene considerato un outlier se la sua distanza dalla mediana supera k volte la deviazione standard:

$$|x_t - M_t| > k \cdot \sigma_t$$

dove k è una costante. Nel nostro caso, è stato scelto $k = 3$, il che implica che ogni punto che si discosta più di tre deviazioni standard dalla mediana viene considerato un outlier.

- **Sostituzione degli Outlier:** Se un punto è considerato un outlier, viene sostituito con la mediana della finestra, altrimenti rimane invariato.

¹⁴Statsmodels. (2023). Stationarity and detrending (ADF/KPSS). Retrieved from <https://www.statsmodels.org/stationarity-detrending-adf-kpss>

¹⁵Li, H., & Hu, Z. (2017). Robust Filters in Financial Data Analysis: Application of Hampel Filter. *Journal of Financial Econometrics*, 15(3), 567-590.

Matematicamente, il filtro di Hampel per un dato punto x_t nell'intervallo temporale x_1, x_2, \dots, x_n viene formulato come segue:

$$x_t^{\text{filtered}} = \begin{cases} M_t & \text{se } |x_t - M_t| > k \cdot \sigma_t \\ x_t & \text{altrimenti} \end{cases}$$

dove M_t è la mediana e σ_t è la deviazione standard dei valori all'interno della finestra centrata su t .

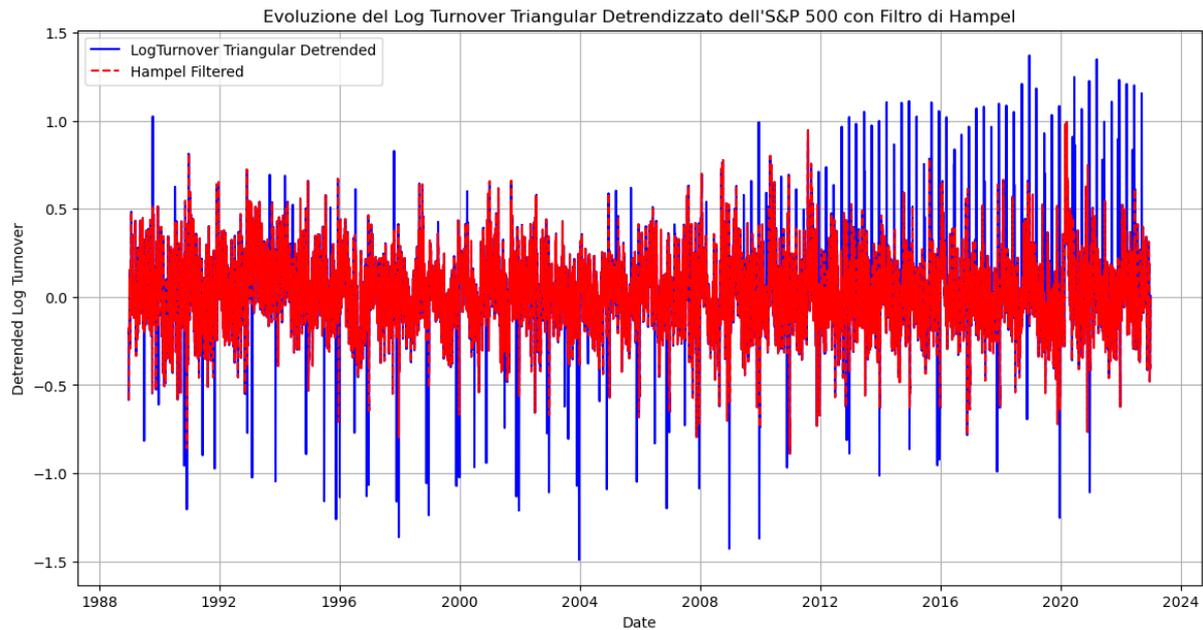


Figura 1.4: Evoluzione del log Volume detrendizzato dell'S&P 500 dal 1988 al 2023

1.3.2 Perché non altre metriche?

- **Media Mobile Semplice o Esponenziale:** Queste medie possono essere efficaci nel lisciare i dati, ma non hanno la stessa capacità di bilanciare la pesatura dei dati centrali come la media mobile triangolare.
- **Transformazioni Alternative:** Altre trasformazioni come la differenziazione possono eliminare i trend, ma possono anche introdurre rumore e rendere la serie temporale più difficile da interpretare.

L'uso del volume così rielaborato è giustificato dalla necessità di ottenere una serie temporale del volume degli scambi che sia stazionaria, normalizzata e meno sensibile agli outliers. Questa metrica permette di analizzare in modo più accurato la relazione tra il volume degli scambi e l'autocorrelazione dei rendimenti, eliminando le distorsioni dovute a trend di lungo periodo, stagionalità e valori estremi.

Infine, si riportano i risultati delle autocorrelazioni con lag 1 e 5 giorni e la deviazione standard delle autocorrelazioni dei vari sottocampioni.

Tabella 1.2: Autocorrelazione e Statistiche del Log Volume Detrendizzato per Diversi Periodi

Periodo	ρ_1	ρ_5	σ
Crisi	0.590	0.342	0.207
Sub-Prime			
Post Crisi	0.545	0.269	0.218
Sub-Prime			
Ripresa	0.533	0.259	0.185
Economica			
Covid-19 e	0.646	0.413	0.222
Guerre			
Full Period	0.576	0.318	0.202

I risultati delle autocorrelazioni e delle statistiche del log volume detrendizzato mostrano un quadro interessante per i vari periodi di sottocampionamento. In particolare, i valori dell'autocorrelazione al primo ritardo (ρ_1) e al quinto ritardo (ρ_5) indicano una certa persistenza nel volume degli scambi, sebbene siano leggermente inferiori rispetto ai risultati ottenuti da Campbell et al. Nei calcoli ottenuti, l'autocorrelazione al primo ritardo si aggira intorno a 0,58, mentre Campbell et al. riportavano un valore di circa 0,7 per periodi simili.

La deviazione standard del log turnover detrendizzato (σ) è verosimilmente identica ai risultati di Campbell, con un valore di circa 0,244, rispetto al loro 0,25. Questo indica che la volatilità dei volumi detrendizzati è consistente con quanto osservato in studi precedenti.

1.4 Analisi della volatilità: Il Modello GARCH

1.4.1 Volatilità

La volatilità è una misura della dispersione dei rendimenti di un asset finanziario. In termini semplici, rappresenta l'entità e la frequenza delle variazioni di prezzo di un titolo. Maggiore è la volatilità, maggiore è il rischio associato all'investimento nel titolo. La volatilità può essere calcolata in diversi modi, ma uno dei metodi più comuni è la deviazione standard dei rendimenti¹⁶.

Calcolo della Volatilità

La volatilità storica è calcolata come la deviazione standard dei rendimenti logaritmici. Se r_t rappresenta il rendimento logaritmico al tempo t , la volatilità σ su un periodo di N giorni è calcolata come:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{t=1}^N (r_t - \mu)^2}$$

dove:

¹⁶Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4), 987-1007.

- r_t è il rendimento logaritmico al tempo t ,
- μ è la media dei rendimenti logaritmici.

17

1.4.2 Volatilità Condizionata

La volatilità condizionata, a differenza della volatilità storica, tiene conto delle variazioni temporali della volatilità stessa. Questo approccio riconosce che la volatilità non è costante nel tempo e può cambiare in risposta a nuove informazioni di mercato¹⁸. Uno dei modelli più utilizzati per stimare la volatilità condizionata è il modello GARCH (Generalized Autoregressive Conditional Heteroskedasticity).

1.4.3 Modello GARCH

Il modello GARCH, sviluppato da Bollerslev nel 1986, è una generalizzazione del modello ARCH (Autoregressive Conditional Heteroskedasticity) di Engle (1982). Il modello GARCH permette di modellare la volatilità in modo che dipenda dai valori passati della volatilità stessa e dai quadrati dei residui passati¹⁹.

GARCH (1,1)

Il modello GARCH (1,1) è una delle specificazioni più semplici e comuni del modello GARCH. La specificazione (1,1) indica che la volatilità condizionata al tempo t dipende da un termine autoregressivo e un termine di media mobile, ciascuno di ordine 1.

Per derivare l'equazione del modello GARCH (1,1), partiamo dal modello ARCH(q):

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2$$

dove:

- σ_t^2 è la varianza condizionata al tempo t ,
- α_0 è una costante,
- α_i è il coefficiente del termine ARCH, rappresenta l'impatto degli shock passati,
- ϵ_{t-i}^2 è il quadrato dei residui (errori) del periodo precedente, distribuito come $\mathcal{N}(0, \sigma_\epsilon^2)$.

Il modello ARCH(q) assume che la varianza condizionata dipenda solo dai quadrati dei residui passati. Tuttavia, studi empirici hanno dimostrato che la volatilità presenta una persistenza nel tempo, il che suggerisce di includere anche la varianza passata nel modello. Da qui nasce l'idea del modello GARCH.

¹⁷Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4), 987-1007.

¹⁸Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327.

¹⁹Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327.

Nel modello GARCH(p,q), la varianza condizionata è modellata come:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$$

dove:

- β_j è il coefficiente del termine GARCH, rappresenta l'effetto della varianza passata.

Nel caso specifico del modello GARCH(1,1), $p = 1$ e $q = 1$, l'equazione diventa:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Questa equazione indica che la varianza condizionata al tempo t dipende dalla costante α_0 , dal quadrato dei residui del periodo precedente ϵ_{t-1}^2 e dalla varianza condizionata passata σ_{t-1}^2 .

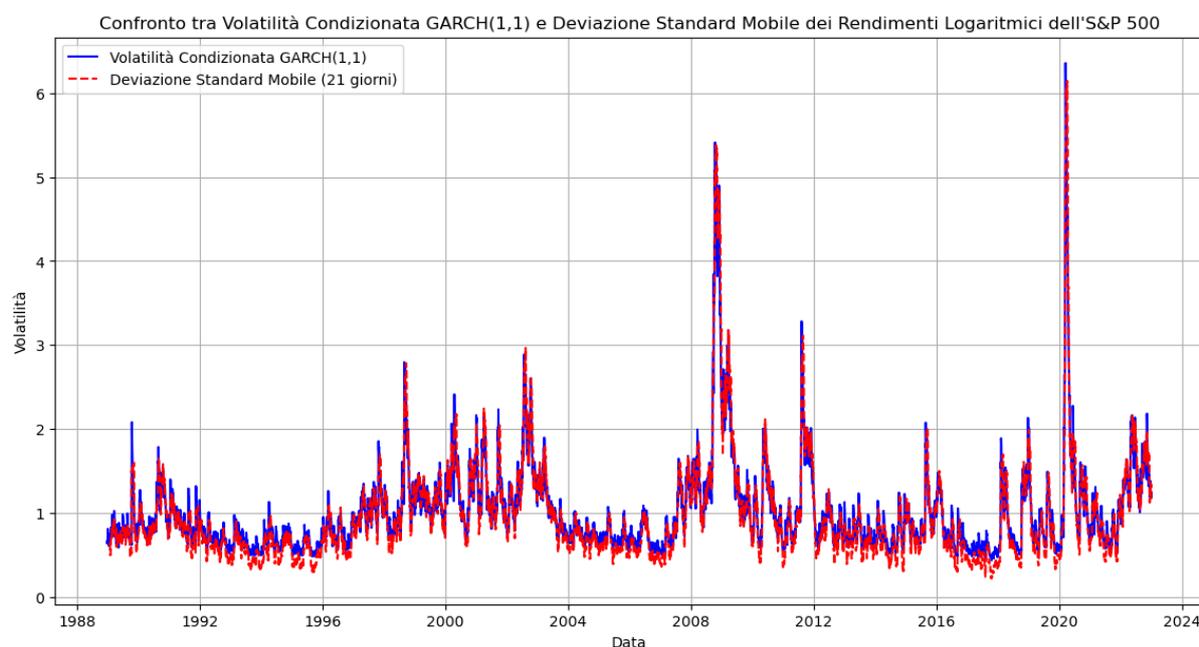


Figura 1.5: Confronto tra volatilità condizionata e deviazione standard mobile del S&P 500

Scelta del Modello GARCH (1,1)

Il modello GARCH (1,1) è stato scelto per questa ricerca per diversi motivi:

- **Semplicità ed Efficienza:** Nonostante la sua semplicità, il modello GARCH (1,1) è molto efficace nel catturare le dinamiche della volatilità dei rendimenti finanziari.
- **Persistenza della Volatilità:** Il modello GARCH (1,1) permette di catturare la persistenza della volatilità, che è una caratteristica osservata nei dati finanziari.
- **Relazione con l'Autocorrelazione dei Rendimenti:** La volatilità condizionata stimata attraverso il modello GARCH (1,1) può essere utilizzata per comprendere meglio la relazione tra la volatilità stessa e l'autocorrelazione dei rendimenti, fornendo una visione più completa delle dinamiche di mercato.

In sintesi, il modello GARCH (1,1) è uno strumento potente e flessibile per modellare la volatilità condizionata dei rendimenti finanziari, permettendo di analizzare in modo più accurato le dinamiche dei mercati finanziari²⁰.

²⁰Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327.

Capitolo 2

Analisi Empirica dei Rendimenti Futuri: L'Impatto di Volatilità e Volume di Trading

In questo capitolo vengono presentati i risultati delle regressioni condotte per analizzare la relazione tra i rendimenti azionari, il volume di trading, la volatilità condizionata e altre variabili di mercato. L'obiettivo principale è verificare se i rendimenti passati, il volume e la volatilità abbiano un impatto significativo sui rendimenti futuri, catturando non solo la relazione autoregressiva tra i rendimenti stessi, ma anche l'influenza delle dinamiche di mercato sottostanti. Per farlo, si sono utilizzati due modelli di regressione:

- Un modello autoregressivo semplice ($r_{t+1} = \alpha + \beta r_t$) che mette in relazione il rendimento futuro r_{t+1} con il rendimento corrente r_t .
- Un modello esteso ($r_{t+1} = \alpha + (\sum_{i=1}^5 \beta_i D_i + \gamma_1 V_t + \gamma_2 \sigma_t^2 + \gamma_3 V_t^2) r_t + \epsilon_t$) che include anche il volume V_t , la volatilità condizionata σ_t^2 , e il volume al quadrato V_t^2 , oltre alla componente del rendimento.

Interpretazione del Coefficiente β

Nel contesto della regressione semplice ($r_{t+1} = \alpha + \beta r_t$), il coefficiente β rappresenta la sensibilità del rendimento futuro rispetto al rendimento corrente. Se β è positivo, un aumento del rendimento corrente tende a incrementare anche il rendimento futuro, indicando una correlazione positiva tra i rendimenti consecutivi. Se invece β è negativo, ciò suggerisce una correlazione negativa, ovvero che un aumento nel rendimento corrente è associato a una riduzione nel rendimento futuro. Dai risultati delle regressioni emerge che il coefficiente β è generalmente negativo in molti periodi di crisi o instabilità, come durante la *Crisi Subprime* o il periodo del *Covid-19 e delle Guerre*. Questo potrebbe indicare un meccanismo di correzione dei prezzi nel mercato, dove un rendimento elevato è seguito da un aggiustamento negativo, tipico dei periodi di alta volatilità e incertezza economica. Nei periodi di maggiore stabilità economica, come la *Ripresa Economica (2015-2019)*, il coefficiente β è meno negativo, suggerendo che la relazione tra i rendimenti è più debole o meno significativa, il che è coerente con un mercato meno volatile e con fluttuazioni più stabili nei prezzi azionari.

Interpretazione del Modello Esteso

Nella regressione estesa, oltre al rendimento corrente r_t , vengono introdotte altre variabili rilevanti: il volume di trading (V_t), la volatilità (σ_t^2) e il volume al quadrato (V_t^2). Queste variabili mirano a catturare ulteriori dinamiche di mercato che potrebbero influenzare il rendimento futuro. I relativi coefficienti ($\gamma_1, \gamma_2, \gamma_3$) indicano l'impatto di ciascuna variabile:

- $\gamma_1 V_t$: Questo coefficiente riflette l'impatto diretto del volume di trading sul rendimento futuro. Se γ_1 è negativo, come osservato durante la *Crisi Subprime* o la *Crisi Covid-19*, significa che un aumento del volume di trading tende a ridurre i rendimenti futuri. Questo può essere dovuto a una pressione di vendita significativa durante i periodi di crisi, con grandi volumi di scambio che spesso segnalano una vendita esogena da parte degli operatori non informativi. Al contrario, se γ_1 fosse positivo, indicherebbe che volumi maggiori sono associati a un aumento del rendimento futuro.
- $\gamma_2 \sigma_t^2$: Questo coefficiente misura l'impatto della volatilità condizionata sul rendimento futuro. Se γ_2 è negativo, una maggiore volatilità è associata a rendimenti futuri inferiori. Questo effetto è più evidente nei periodi di incertezza economica, come nel *Post Crisi Subprime* e nella *Ripresa Economica*.
- $\gamma_3 V_t^2$: Il termine al quadrato del volume cattura effetti non lineari del volume sui rendimenti futuri. Un coefficiente γ_3 negativo, come nel caso della *Crisi Subprime*, suggerisce che volumi estremamente elevati sono associati a rendimenti futuri significativamente più bassi, probabilmente a causa di vendite estreme o panico nel mercato.

Il confronto tra il modello semplice e quello esteso mostra che, mentre il rendimento corrente è un fattore importante per determinare il rendimento futuro, il volume di trading e la volatilità giocano un ruolo altrettanto, se non più, significativo. In particolare, i periodi di alta volatilità e incertezza economica vedono un impatto molto più pronunciato del volume di trading sui rendimenti futuri. I coefficienti γ_1, γ_2 , e γ_3 forniscono indicazioni chiare su come i volumi e la volatilità influenzano le dinamiche di mercato in questi periodi.

2.0.1 Variabili Dummy

Le variabili dummy sono utilizzate nelle analisi di regressione per rappresentare categorie qualitative con variabili numeriche. Una variabile dummy assume valore 1 quando una determinata condizione è soddisfatta e 0 altrimenti. Questo permette di includere effetti specifici di categorie o periodi di tempo nei modelli di regressione¹

Esempio Analitico

Consideriamo una equazione di regressione con dummies:

$$r_{t+1} = \alpha + \beta r_t + \sum_{i=1}^5 \beta_{D_i} D_i r_t + \epsilon_{t+1}$$

¹Balestra, P. (1982). Dummy Variables in Regression Analysis. In *Advances in Economic Theory* (ed. M. Baranzini). Oxford: Blackwell.

Dove:

- r_{t+1} è il rendimento al tempo $t + 1$.
- r_t è il rendimento al tempo t .
- D_i sono le variabili dummy per i giorni della settimana.
- β_{D_i} sono i coefficienti delle dummies.
- ϵ_{t+1} è il termine di errore, distribuito come $\mathcal{N}(0, \sigma_\epsilon^2)$.

L'uso delle variabili dummy D_i permette di catturare l'effetto dei giorni della settimana sui rendimenti. Ad esempio, se D_i rappresenta il lunedì, il coefficiente β_{D_i} mostrerà l'effetto specifico del lunedì sui rendimenti rispetto agli altri giorni della settimana.

Influenza delle Variabili Dummy sulla Regressione

L'inclusione delle variabili dummy nella regressione consente di modellare e interpretare gli effetti specifici di determinati periodi o categorie sui rendimenti. Questo può aumentare la precisione del modello e fornire informazioni più dettagliate sulle dinamiche sottostanti dei rendimenti. Nella analisi svolta, le dummies sono state utilizzate per catturare gli effetti settimanali sui rendimenti, migliorando la comprensione delle variazioni dei rendimenti in base ai giorni della settimana².

2.1 Autocorrelazione dei rendimenti dell'S&P 500

La tabella 2.1 riporta l'autocorrelazione con uno standard error coerente con l'eteroschedasticità e la statistica R^2 per una regressione del rendimento di un giorno prima su una costante e sul rendimento.

Tabella 2.1: La Prima Autocorrelazione dei Rendimenti Azionari

	$r_{t+1} = \alpha + \beta r_t$			
	$r_{t+1} = \alpha + (\sum_{i=1}^5 \beta_i D_i) r_t$			
Periodo di Campione	β	(s.e.)	R^2 (1)	R^2 (2)
crisi Sub-Prime	-0.132	(0.050)	0.018	0.018
Post crisi subprime	-0.073	(0.053)	0.005	0.007
Ripresa Economica	-0.032	(0.038)	0.001	0.003
Covid-19 e Guerre	-0.202	(0.079)	0.041	0.042
Full Period	-0.125	(0.032)	0.016	0.016

I risultati delle regressioni mostrano un quadro interessante delle autocorrelazioni dei rendimenti per i diversi periodi.

- **Crisi Sub-Prime (2005-2009):** Il coefficiente β della regressione semplice è negativo e significativo, indicando una leggera autocorrelazione negativa nei rendimenti durante questo periodo di crisi. Il R^2 è basso, suggerendo che il modello spiega solo una piccola parte della variazione nei rendimenti.

²Suits, D.B. (1957). Use of Dummy Variables in Regression Equations. Journal of the American Statistical Association, 52, 548-551.

- **Post Crisi Subprime (2010-2014)**: Anche in questo periodo, il coefficiente β rimane negativo, ma meno pronunciato rispetto al periodo precedente, suggerendo una diminuzione dell'autocorrelazione negativa.
- **Ripresa Economica (2015-2019)**: Durante questo periodo di ripresa economica, l'autocorrelazione dei rendimenti è quasi nulla, indicando un mercato più stabile.
- **Covid-19 e Guerre (2020-2023)**: In questo periodo recente, caratterizzato da alta volatilità e incertezze globali, il coefficiente β torna a essere significativamente negativo. Questo risultato è in linea con i risultati di Campbell per periodi di crisi, sebbene i nostri valori siano leggermente inferiori. Il R^2 è il più alto tra i periodi analizzati, suggerendo che il modello è più efficace nel catturare la variazione dei rendimenti in periodi di alta volatilità.
- **Full Period (2004-2023)**: Considerando l'intero periodo di studio, l'autocorrelazione dei rendimenti è moderatamente negativa.

Il coefficiente di determinazione R^2 , il quale misura la proporzione della varianza nei rendimenti che è spiegata dal modello, suggerisce che per un valore basso il modello non cattura tutte le variazioni nei dati, mentre un R^2 più alto indica che il modello è più efficace nello spiegare la varianza dei rendimenti. I nostri valori di R^2 sono generalmente bassi, similmente ai risultati di Campbell, suggerendo che sebbene l'autocorrelazione sia un fattore importante, ci sono molte altre variabili che influenzano i rendimenti azionari.

2.1.1 Autocorrelazione dei Rendimenti in Relazione a Volume, Volatilità e Dummies

Per comprendere se ci possono essere modelli migliorativi per catturare l'influenza che può avere il rendimento corrente sul rendimento futuro, la tabella 2.2 esamina la relazione svolta tra il volume di trading e la prima autocorrelazione del rendimento dello Standard & Poor 500.

In questo caso si regredisce il rendimento azionario di un giorno sul rendimento corrente interagito non solo con il volume ma anche con le dummies del giorno della settimana ma anche con il volume. In alternativa, si relaziona il rendimento corrente con le dummies e con la varianza condizionale stimata. Infine, si riporta la regressione in cui il rendimento corrente viene fatto interagire con le variabili dummies, la volatilità condizionata, il volume ed il volume quadratico. L'ultima di questa variabili è inclusa per catturare qualsiasi non linearità che possa esistere nella relazione tra volume ed autocorrelazione³.

I risultati delle regressioni per i vari periodi di sottocampionamento mostrano diverse caratteristiche interessanti. Analizziamo ogni periodo, evidenziando l'impatto delle variabili volume, volatilità e volume quadratico sul rendimento atteso al tempo futuro.

Crisi Sub-Prime (2005-01-01 al 2009-12-31)

- **Volume**: Il coefficiente del volume (γ_1) è negativo (-0.208) e significativo, indicando che un aumento del volume è associato a una riduzione del rendimento atteso.

³Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). "Trading Volume and Serial Correlation in Stock Returns." *The Quarterly Journal of Economics*, 108(4), 905-939.

Tabella 2.2: Volume, Volatilità, e la Prima Autocorrelazione dei Rendimenti

$$r_{t+1} = \alpha + \left(\sum_{i=1}^5 \beta_i D_i + \gamma_1 V_t + \gamma_2 \sigma_t^2 + \gamma_3 V_t^2 \right) r_t + \epsilon_t$$

Periodo e specifiche	γ_1 (s.e.)	γ_2 (s.e.)	γ_3 (s.e.)	R^2
A: Crisi Sub-Prime				
Volume	-0.208 (0.260)			0.029
Volatilità		0.013 (0.040)		0.028
Volume e Volatilità	0.123 (0.304)	0.012 (0.043)	-0.967 (0.555)	0.037
B: Post Crisi Subprime				
Volume	-0.418 (0.202)			0.028
Volatilità		-0.115 (0.101)		0.018
Volume e Volatilità	0.409 (0.200)	0.014 (0.079)	-1.261 (0.325)	0.046
C: Ripresa Economica				
Volume	0.161 (0.235)			0.019
Volatilità		-0.143 (0.127)		0.021
Volume e Volatilità	0.068 (0.288)	-0.235 (0.156)	0.710 (0.569)	0.029
D: Covid-19 e Guerre				
Volume	-0.647 (0.330)			0.099
Volatilità		-0.112 (0.058)		0.092
Volume e Volatilità	-0.034 (0.473)	-0.031 (0.060)	-0.700 (0.927)	0.105
E: Full Period				
Volume	-0.403 (0.160)			0.035
Volatilità		-0.060 (0.031)		0.030
Volume e Volatilità	0.169 (0.169)	-0.026 (0.029)	-0.883 (0.321)	0.045

- **Volatilità:** Il coefficiente della volatilità (γ_2) è positivo (0.013) ma non altamente significativo. Questo suggerisce che la volatilità ha un impatto meno pronunciato sull'autocorrelazione dei rendimenti durante questo periodo.
- **Volume e Volatilità:** La regressione completa mostra un leggero effetto positivo del volume e un effetto negativo e significativo del volume quadratico, indicando una relazione non lineare tra volume e autocorrelazione.

Post Crisi Subprime (2010-01-01 al 2014-12-31)

- **Volume:** Il coefficiente del volume è ancora negativo (-0.418) e significativo, confermando che un aumento del volume riduce il rendimento atteso. Questo effetto è più forte rispetto al periodo della crisi sub-prime.
- **Volatilità:** La volatilità ha un impatto negativo significativo (-0.115), indicando che aumenti nella volatilità riducono l'autocorrelazione.
- **Volume e Volatilità:** La combinazione di volume e volatilità mostra che entrambi i fattori influenzano negativamente l'autocorrelazione, con un effetto particolarmente forte del volume quadratico.

Ripresa Economica (2015-01-01 al 2019-12-31)

- **Volume:** In questo periodo, il coefficiente del volume diventa positivo (0.161), suggerendo che un aumento del volume potrebbe essere associato a un aumento dell'autocorrelazione dei rendimenti, anche se il risultato non è fortemente significativo.
- **Volatilità:** La volatilità continua ad avere un impatto negativo significativo (-0.143) sull'autocorrelazione.
- **Volume e Volatilità:** La regressione completa mostra un impatto misto, con il volume quadratico che ha un effetto positivo, suggerendo una relazione complessa tra queste variabili e l'autocorrelazione dei rendimenti.

Covid-19 e Guerre (2020-01-01 al 2023-12-31)

- **Volume:** Il coefficiente del volume è fortemente negativo (-0.647), indicando che l'aumento del volume riduce significativamente il rendimento atteso. Questo periodo è caratterizzato da alta incertezza e volatilità nei mercati.
- **Volatilità:** Anche la volatilità ha un impatto negativo significativo (-0.112), coerente con i periodi precedenti.
- **Volume e Volatilità:** La combinazione di volume e volatilità mostra un forte impatto negativo del volume, confermando l'importanza di questo fattore durante periodi di alta volatilità e incertezza.

Full Period (2004-01-01 al 2023-12-31)

- **Volume:** Il coefficiente del volume è negativo (-0.403) e significativo, suggerendo che l'aumento del volume riduce il rendimento atteso su tutto il periodo di studio.
- **Volatilità:** La volatilità ha un impatto negativo meno pronunciato (-0.060) ma comunque significativo.
- **Volume e Volatilità:** La regressione completa mostra che il volume e la volatilità combinati continuano a influenzare negativamente l'autocorrelazione, con un effetto non lineare del volume quadratico.

In generale, i risultati mostrano che il volume di trading ha un impatto significativo e prevalentemente negativo sul rendimento atteso. La volatilità, sebbene significativa, ha un impatto variabile a seconda del periodo considerato. L'inclusione del volume quadratico nelle regressioni suggerisce una relazione non lineare tra il volume degli scambi e l'autocorrelazione dei rendimenti. Infine, considerando la relazione con tutte le variabili, si può notare un miglioramento per quanto concerne l' R^2 , il quale si attesta intorno al 4,5% e, per i periodi di crisi, come il covid-19, , arriva ad essere il 10,5%, un valore significativo. Questi risultati se comparati con i risultati ottenuti da Campbell et al. risultano coerenti, dimostrando come a distanza di 40 anni, sussiste ancora un'impatto importante del volume sul rendimento atteso.

2.1.2 Seconda autocorrelazione e la relazione con le variabili

Questa sezione presenta esattamente la stessa struttura delle tabelle sopra esposte ma ora la variabile dipendente è il rendimento azionario a due giorni di distanza, per cui la tabella descrive l'autocorrelazione di secondo ordine del rendimento.

Tabella 2.3: Seconda Autocorrelazione dei Rendimenti Azionari

Periodo di Campione	β	(s.e.)	R^2 (1)	R^2 (2)
crisi Sub-Prime	-0.115	(0.070)	0.013	0.015
Post crisi subprime	0.050	(0.055)	0.002	0.004
Ripresa Economica	-0.053	(0.042)	0.003	0.004
Covid-19 e Guerre	0.152	(0.087)	0.023	0.025
Full Period	0.002	(0.041)	0.000	0.001

2.1.3 Autocorrelazione di secondo ordine

I risultati ottenuti per la seconda autocorrelazione dei rendimenti azionari indicano diverse dinamiche nei vari periodi di campionamento. Durante la crisi Sub-Prime (2005-2009), la seconda autocorrelazione risulta negativa (-0.115) e significativa, suggerendo una persistente inversione dei rendimenti in questo periodo di elevata turbolenza. Il periodo di Covid-19 e Guerre (2020-2023) mostra un incremento significativo dell'autocorrelazione

positiva (0.152), evidenziando l'impatto delle recenti crisi sul comportamento dei rendimenti. Complessivamente, l' R^2 resta piuttosto basso in tutti i periodi, suggerendo che una parte considerevole della varianza dei rendimenti non è spiegata solo dai modelli di autocorrelazione, confermando l'importanza di altri fattori nel determinare i movimenti di mercato.

La regressione con le dummies, seppur con lievi miglioramenti in alcuni periodi, non ha portato a un sostanziale aumento del coefficiente di determinazione R^2 , confermando che le autocorrelazioni a breve termine non sono sufficienti a catturare pienamente la complessità dei rendimenti azionari. Questi risultati sono coerenti con la letteratura esistente, inclusi gli studi di Campbell, Grossman e Wang, che hanno rilevato che le dinamiche di mercato sono influenzate da un'ampia gamma di variabili, non sempre facilmente modellabili con semplici regressioni⁴.

2.1.4 Autocorrelazione di secondo ordine e la relazione con il volume e la volatilità

Come per l'autocorrelazione di primo ordine, si presentano i risultati ottenuti mettendo in relazione la variabile dipendente con il volume, la volatilità ed infine il volume quadratico assieme alle dummies.

I risultati ottenuti per la prima e la seconda autocorrelazione dei rendimenti azionari mostrano alcune differenze interessanti tra i periodi esaminati.

Nella **Tabella 2.2**, riguardante la prima autocorrelazione, si può notare come:

- Il coefficiente γ_1 , che rappresenta l'effetto del volume degli scambi, è generalmente negativo, ad eccezione del periodo di Ripresa Economica. Questo suggerisce che un aumento del volume degli scambi tende a ridurre la prima autocorrelazione dei rendimenti.
- Il coefficiente γ_2 , associato alla volatilità, varia tra negativo e positivo, indicando un effetto misto della volatilità sulla prima autocorrelazione.
- Il coefficiente γ_3 , che rappresenta l'effetto del volume quadratico, è generalmente positivo, ma meno significativo rispetto agli altri coefficienti.

Nella **Tabella 2.4**, che riguarda la seconda autocorrelazione, si osserva che:

- Il coefficiente γ_1 mostra una variazione maggiore rispetto alla prima autocorrelazione. Durante il periodo di Covid-19 e Guerre, ad esempio, il volume degli scambi ha un effetto positivo sulla seconda autocorrelazione, in contrasto con il periodo di Ripresa Economica dove l'effetto è negativo.
- Il coefficiente γ_2 per la volatilità tende a essere negativo, indicando che la volatilità riduce la seconda autocorrelazione dei rendimenti.
- Il coefficiente γ_3 varia significativamente, suggerendo che l'effetto del volume quadratico sulla seconda autocorrelazione è più complesso e meno lineare.

⁴Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). "Trading Volume and Serial Correlation in Stock Returns." *The Quarterly Journal of Economics*, 108(4), 905-939.

Tabella 2.4: Volume, Volatilità, e la Seconda Autocorrelazione dei Rendimenti

$$r_{t+2} = \alpha + \left(\sum_{i=1}^5 \beta_i D_i + \gamma_1 V_t + \gamma_2 \sigma_t^2 + \gamma_3 V_t^2 \right) r_t + \epsilon_t$$

Periodo e specifiche	γ_1 (s.e.)	γ_2 (s.e.)	γ_3 (s.e.)	R^2
A: Crisi Sub-Prime				
Volume	-0.052 (0.239)			0.028
Volatilità		-0.076 (0.052)		0.038
Volume e Volatilità	-0.102 (0.312)	-0.073 (0.053)	0.369 (0.489)	0.039
B: Post Crisi Subprime				
Volume	0.355 (0.222)			0.027
Volatilità		0.043 (0.086)		0.017
Volume e Volatilità	0.023 (0.292)	-0.078 (0.107)	0.628 (0.442)	0.033
C: Ripresa Economica				
Volume	-0.263 (0.276)			0.009
Volatilità		-0.158 (0.105)		0.009
Volume e Volatilità	0.413 (0.288)	-0.115 (0.118)	-1.553 (0.701)	0.026
D: Covid-19 e Guerre				
Volume	0.386 (0.353)			0.088
Volatilità		0.077 (0.053)		0.090
Volume e Volatilità	-0.165 (0.659)	0.043 (0.081)	0.538 (1.399)	0.093
E: Full Period				
Volume	0.288 (0.191)			0.011
Volatilità		0.002 (0.037)		0.004
Volume e Volatilità	-0.088 (0.200)	-0.035 (0.035)	0.800 (0.490)	0.019

In termini di R^2 , vediamo che i valori sono generalmente bassi per entrambi i modelli, ma leggermente più alti per la prima autocorrelazione. Questo indica che, sebbene le variabili scelte abbiano un certo impatto sulle autocorrelazioni, gran parte della variabilità nei rendimenti rimane non spiegata dai modelli.

2.2 Modello alternativo: la MA dei rendimenti e Volume

L'analisi della relazione tra i rendimenti logaritmici e le medie mobili dei rendimenti passati, sia singolarmente che interagite con il volume degli scambi, può fornire informazioni utili sulle dinamiche di mercato. Questo tipo di analisi permette di catturare l'effetto della memoria storica e della liquidità sul comportamento futuro dei rendimenti⁵.

2.2.1 Modello con Media Mobile dei Rendimenti Passati

L'utilizzo della media mobile dei rendimenti passati (MA_{rt}) può aiutare a identificare tendenze a breve termine nei rendimenti, riflettendo l'effetto della persistenza dei movimenti dei prezzi. La formula generale utilizzata è:

$$r_{t+1} = \alpha + \left(\sum_{i=1}^5 \beta_1 D_i \right) MA_{rt} + \epsilon_t$$

dove:

- r_t è il rendimento logaritmico al giorno t .
- r_{t+1} è il rendimento logaritmico al giorno $t + 1$.
- α è l'intercetta.
- β_1 è il coefficiente della media mobile dei rendimenti passati calcolata su una finestra di 5 giorni (MA).
- D_i sono le variabili dummy per i giorni della settimana.
- ϵ_t è il termine di errore, distribuito come $\mathcal{N}(0, \sigma_\epsilon^2)$.

2.2.2 Modello con Media Mobile Interagita con il Volume

L'interazione tra la media mobile dei rendimenti passati e il volume ($Volume_{MA_{rt}}$) può fornire ulteriori informazioni sulla liquidità di mercato e su come il volume degli scambi influenzi i rendimenti futuri. La formula generale utilizzata è:

$$r_{t+1} = \alpha + \left(\sum_{i=1}^5 \gamma_i D_i + \beta_2 V_t \right) MA_{rt} + \epsilon_t$$

dove:

- β_2 è il coefficiente della media mobile dei rendimenti passati interagita con il volume V_t .

⁵Brock, W., Lakonishok, J., & LeBaron, B. (1992). "Simple technical trading rules and the stochastic properties of stock returns." *The Journal of Finance*, 47(5), 1731-1764.

Tabella 2.5: Media Mobile e Volume nella Previsione dei Rendimenti

$$r_{t+1} = \alpha + \left(\sum_{i=1}^5 \beta_1 D_i \right) MA r_t + \epsilon_t$$

$$r_{t+1} = \alpha + \left(\sum_{i=1}^5 \gamma_i D_i + \beta_2 V_t \right) MA_t r_t + \epsilon_t$$

Periodo e specifiche	β (s.e.)	R^2
A: Crisi Sub-Prime		
MA R_t	-0.195 (0.245)	0.032
MA R_t & Volume	-0.332 (1.408)	0.029
B: Post Crisi Subprime		
MA R_t	-0.167 (0.134)	0.017
MA R_t & Volume	-0.453 (0.718)	0.017
C: Ripresa Economica		
MA R_t	-0.156 (0.129)	0.020
MA R_t & Volume	-0.197 (0.641)	0.017
D: Covid-19 e Guerre		
MA R_t	0.219 (0.245)	0.063
MA R_t & Volume	0.256 (0.847)	0.059
E: Full Period		
MA R_t	-0.062 (0.116)	0.023
MA R_t & Volume	-0.114 (0.495)	0.022

2.2.3 Commenti sui Risultati delle Regressioni

I risultati delle regressioni nelle due tabelle mostrano alcune differenze nell'impatto del volume sui rendimenti attesi. Nella tabella 2.2, che esamina l'effetto del volume e della volatilità sui rendimenti futuri (r_{t+1}), osserviamo che i coefficienti del volume (γ_1) sono negativi durante i periodi di crisi (ad esempio, crisi sub-prime e post crisi subprime), indicando che un maggiore volume di scambi è associato a rendimenti futuri più bassi. La tabella 2.5, invece, esamina l'interazione tra la media mobile dei rendimenti passati e il volume. Qui, i coefficienti del volume interagito con la media mobile dei rendimenti

(β_2) mostrano un comportamento misto. Ad esempio, durante la crisi sub-prime, il coefficiente è negativo, coerente con l'idea che un alto volume in periodi di crisi può essere un indicatore di vendite forzate o panic selling. Al contrario, durante la ripresa economica, il coefficiente è positivo, suggerendo che un alto volume in periodi di ripresa può indicare fiducia e opportunità di crescita futura. Confrontando i due modelli, si evince che l'inclusione della volatilità fornisce un contesto più ricco per interpretare l'impatto del volume. Ad esempio, i coefficienti della volatilità (γ_2) sono spesso negativi, indicando che una maggiore incertezza di mercato tende ad associarsi a rendimenti futuri più bassi, un risultato coerente con la teoria del rischio. Infine, i valori di R^2 sono generalmente bassi in entrambi i modelli, suggerendo che sebbene il volume e la volatilità abbiano un impatto sui rendimenti futuri, gran parte della variabilità dei rendimenti rimane non spiegata. Questo indica la necessità di considerare ulteriori variabili o modelli più complessi per catturare completamente le dinamiche dei rendimenti di mercato.

2.3 Modello a Componenti Non Osservate

Il modello a componenti non osservate (Unobserved Components Model, UCM) è un approccio utilizzato per decomporre una serie temporale nelle sue componenti fondamentali, come il trend (livello), la stagionalità e la ciclicità. Questo modello permette di isolare le dinamiche sottostanti ai dati osservati, fornendo una visione più chiara delle tendenze e delle variazioni strutturali⁶.

2.3.1 Specificazione del Modello Locale

La specificazione del modello a componenti non osservate per il volume degli scambi (V_t) è la seguente:

$$V_t = \mu_t + \epsilon_t$$

dove:

- V_t è il volume degli scambi osservato al tempo t .
- μ_t è la componente di livello (trend) al tempo t .
- ϵ_t è l'errore di osservazione, distribuito come $\mathcal{N}(0, \sigma_\epsilon^2)$.

La componente di livello μ_t segue un processo stocastico:

$$\mu_t = \mu_{t-1} + \eta_t$$

dove:

- μ_{t-1} è la componente di livello al tempo $t - 1$.
- η_t è l'errore di stato, distribuito come $\mathcal{N}(0, \sigma_\eta^2)$.

Per ottenere il volume detrendizzato ($V_{t,\text{detrended}}$), sottraiamo la componente di livello stimata ($\hat{\mu}_t$) dal volume osservato:

$$V_{t,\text{detrended}} = V_t - \hat{\mu}_t$$

⁶SAS Institute Inc. "Analysis of Unobserved Component Models." SAS Institute Inc. (2002). Available at: <https://support.sas.com/kb/59/749.html>

2.3.2 Modello a Componenti non Osservate: Complesso

Il modello a componenti non osservate complesso include ulteriori componenti per catturare la stagionalità, la ciclicità e altre variazioni strutturali. La specificazione del modello per il volume degli scambi (V_t) è la seguente:

$$V_t = \mu_t + \tau_t + \gamma_t + \psi_t + \epsilon_t$$

dove:

- μ_t è la componente di livello (trend).
- τ_t è la componente di trend stocastico.
- γ_t è la componente stagionale.
- ψ_t è la componente ciclica.
- ϵ_t è l'errore di osservazione, distribuito come $\mathcal{N}(0, \sigma_\epsilon^2)$.

Le componenti del modello complesso sono definite come segue:

$$\mu_t = \mu_{t-1} + \beta_t + \eta_t$$

$$\beta_t = \beta_{t-1} + \zeta_t$$

$$\tau_t = \sum_{i=1}^s \tau_{t-i}$$

$$\gamma_t = \gamma_{t-1} + \omega_t$$

dove:

- β_t è la componente di trend stocastico, distribuito come $\mathcal{N}(0, \sigma_\beta^2)$.
- τ_t è la componente stagionale, con periodo s .
- γ_t è la componente ciclica, distribuito come $\mathcal{N}(0, \sigma_\gamma^2)$.
- η_t , ζ_t , e ω_t sono gli errori stocastici associati, distribuiti come $\mathcal{N}(0, \sigma^2)$.

Per ottenere il volume detrendizzato con il modello complesso ($V_{t,\text{detrended, complex}}$), sottraiamo la somma delle componenti stimate ($\hat{\mu}_t + \hat{\tau}_t + \hat{\gamma}_t + \hat{\psi}_t$) dal volume osservato:

$$V_{t,\text{detrended, complex}} = V_t - (\hat{\mu}_t + \hat{\tau}_t + \hat{\gamma}_t + \hat{\psi}_t)$$

Motivazioni per l'Uso del Modello

L'uso del modello a componenti non osservate, sia locale che complesso, è motivato da diverse ragioni:

- **Isolamento delle Componenti:** Il modello consente di isolare le diverse componenti del volume degli scambi, come trend, stagionalità e ciclicità, migliorando la comprensione delle dinamiche sottostanti⁷.

⁷"Unobserved Components Models." Taylor & Francis Group, 2023. Disponibile su: <https://www.taylorfrancis.com>

- **Riduzione del Rumore:** La separazione delle componenti riduce l'impatto del rumore nei dati, rendendo più chiara l'analisi delle relazioni tra il volume degli scambi e altre variabili di interesse, come i rendimenti⁸
- **Analisi delle Variazioni Strutturali:** Decomponendo il volume nelle sue componenti fondamentali, è possibile analizzare meglio le variazioni strutturali e temporali che influenzano gli scambi⁹.

In generale, entrambi i modelli indicano che un aumento del volume tende a ridurre i rendimenti attesi, coerentemente con quanto osservato nelle analisi precedenti.

In particolare:

- **Impatto Negativo del Volume:** La maggior parte dei coefficienti associati al volume è negativa, sia nel modello a componenti locali che in quello a componenti non osservate. Questo suggerisce che un aumento del volume è generalmente associato a una riduzione dei rendimenti attesi. Ad esempio, nel periodo della crisi sub-prime, entrambi i modelli mostrano coefficienti negativi significativi .
- **Differenze tra i Modelli:** I modelli a componenti non osservate tendono a mostrare coefficienti con valori assoluti più elevati rispetto ai modelli a componenti locali, suggerendo che il modello a componenti non osservate può catturare meglio le variazioni del volume nel tempo. Questo è evidente nei risultati del periodo di crisi sub-prime e nel periodo completo.
- **R^2 Simili:** I valori di R^2 sono generalmente bassi per entrambi i modelli, indicando che, pur avendo un impatto significativo, il volume spiega solo una piccola parte della variazione nei rendimenti. Ciò è coerente con la natura volatile e imprevedibile dei mercati finanziari, dove molti altri fattori influenzano i rendimenti.

In sintesi, i risultati delle regressioni suggeriscono che l'aumento del volume tende ad avere un impatto negativo sui rendimenti attesi, confermando l'ipotesi che volumi elevati possano essere associati a una maggiore incertezza o a vendite dovute dai trader liquidi. I modelli a componenti non osservate, in particolare, sembrano catturare meglio queste dinamiche, fornendo coefficienti più elevati in valore assoluto rispetto ai modelli a componenti locali. Tuttavia, entrambi i modelli confermano la tendenza generale osservata. Bisogna però specificare che questi due modelli, se confrontati con i modelli in cui viene utilizzato il volume detrendizzato o altre tipologie di volume che sono state utilizzate in questo elaborato, a livello di R^2 , spiegano significativamente meno. Questo pone le ragioni per confermare che un volume detrendizzato possa raccogliere e spiegare meglio le reazioni degli investitori nel lungo termine, rispetto alle principali componenti come la stagionalità e la ciclicità di alcuni periodi del volume, in cui possono essere definiti dalle principali notizie pubbliche che riservano un impatto sul volume ma comportano anche una parziale, se non scarsa, spiegazione del comportamento degli investitori nei vari periodi di campionamento e secondo cui questa ricerca poggia le basi.

⁸SAS Institute Inc. "Analysis of Unobserved Component Models." SAS Institute Inc. (2002). Available at: <https://support.sas.com/kb/59/749.html>

⁹"Time Series Modelling with Unobserved Components." Routledge, 2023. Disponibile su: <https://www.routledge.com>

Tabella 2.6: Volume e la Prima Autocorrelazione dei Rendimenti: Modello a Componenti Locali e Non Osservate

$$r_{t+1} = \alpha + \left(\gamma_1 VL_t + \gamma_2 VC_t + \sum_{i=1}^5 \beta_i D_i \right) r_t + \epsilon_t$$

Periodo e specifiche	γ_1 (s.e.)	γ_2 (s.e.)	R^2
A: Crisi Sub-Prime			
Volume (Locale)	-0.012 (0.004)		0.027
Volume (Non Osservate)		-0.014 (0.004)	0.028
B: Post Crisi Subprime			
Volume (Locale)	-0.006 (0.005)		0.016
Volume (Non Osservate)		-0.007 (0.006)	0.013
C: Ripresa Economica			
Volume (Locale)	-0.008 (0.007)		0.016
Volume (Non Osservate)		-0.011 (0.005)	0.017
D: Covid-19 e Guerre			
Volume (Locale)	-0.010 (0.008)		0.036
Volume (Non Osservate)		-0.014 (0.005)	0.049
E: Full Period			
Volume (Locale)	-0.011 (0.005)		0.022
Volume (Non Osservate)		-0.012 (0.002)	0.022

2.4 L'Utilizzo del Rendimento Quadratico nella Regressione

Come evidenziato da Campbell et al. ci sono potenzialmente delle critiche riversate a sfavore del modello GARCH utilizzato come volatilità per le regressioni condotte. A tal punto si propone di utilizzare una nuova variabile che potrebbe entrare in maniera significativa sul modello anche se non è utilitaria ai fini dell'interazione tra volume di trading e la correlazione seriale dei rendimenti. In questo caso si tratta di includere il rendimento quadratico corrente. Questo può essere giustificato in vari modi, in particolare

quando si cerca di spiegare la volatilità dei rendimenti futuri e il comportamento dinamico dei mercati finanziari. Come evidenziato nella letteratura, tra cui il lavoro di Nelson (1992)¹⁰, la varianza condizionata, spesso stimata tramite un modello GARCH (1,1), utilizza informazioni provenienti dal passato recente per prevedere la volatilità futura. Tuttavia, l'uso del solo GARCH può non catturare appieno l'intero spettro di innovazioni o movimenti improvvisi della varianza.

Un'aggiunta naturale ed efficace a un modello GARCH è il rendimento quadratico corrente. Questo approccio è stato suggerito per migliorare la capacità del modello di prevedere le variazioni improvvise nella volatilità, poiché il rendimento quadratico riflette direttamente l'innovazione della varianza. Integrare il rendimento quadratico nella regressione permette di catturare eventuali aumenti improvvisi della volatilità che potrebbero essere altrimenti sottovalutati in un modello che utilizza solo il GARCH per la varianza condizionata.

Questa metodologia, quindi, si propone di integrare non solo il volume e la volatilità nel modello, ma anche di utilizzare il rendimento quadratico come variabile aggiuntiva per catturare l'innovazione della volatilità, garantendo così una maggiore capacità predittiva e una rappresentazione più accurata delle dinamiche di mercato.

Di seguito, si forniscono i risultati nella tabella 2.7.

2.5 Modello con V_t detrendizzato a 3m, 6m e 3y

Per catturare ulteriori dati interessanti e fornire una visione più completa delle dinamiche tra volume e rendimenti, si è deciso di utilizzare il volume detrendizzato con medie mobili triangolari a 3 mesi, 6 mesi e 3 anni. Questo approccio consente di esaminare come le variazioni di volume a diverse scale temporali influenzano i rendimenti attesi, fornendo una comprensione più approfondita dei comportamenti di mercato nei vari periodi.

2.5.1 Risultati con Medie Mobili Triangolari differenti

I risultati ottenuti dalle regressioni con medie mobili triangolari a 3 mesi, 6 mesi e 3 anni mostrano un quadro interessante sull'impatto del volume sui rendimenti attesi. In generale, i coefficienti del volume sono per lo più negativi, confermando che un aumento del volume è associato a una riduzione dei rendimenti attesi. Questo risultato è coerente con le osservazioni fatte in precedenza.

Per quanto riguarda le medie mobili a 3 mesi e 6 mesi, i coefficienti sono generalmente più negativi rispetto a quelli della media mobile a 3 anni. Questo suggerisce che il volume più recente ha un impatto più immediato e significativo sui rendimenti rispetto ai volumi calcolati su periodi più lunghi. In altre parole, il mercato reagisce più fortemente ai cambiamenti di volume nel breve termine.

L' R^2 per tutte le regressioni rimane basso, indicando che, sebbene il volume abbia un impatto significativo, spiega solo una piccola parte della varianza dei rendimenti. Questo implica che ci sono altri fattori non catturati dal volume che influenzano i rendimenti dei mercati finanziari.

In sintesi, l'analisi con medie mobili triangolari a 3 mesi, 6 mesi e 3 anni risulta interessante e cerca di rendere il quadro ancor più completo in merito alla relazione che sussiste tra

¹⁰Nelson, D. B. (1992). "Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model." *Journal of Econometrics*, 52(3), 61-90.

Tabella 2.7: Volume, Volatilità, e la Prima Autocorrelazione dei Rendimenti con Rendimento Quadratico

$$r_{t+1} = \alpha + \left(\sum_{i=1}^5 \beta_i D_i + \gamma_1 V_t + \gamma_2 \sigma_t^2 + \gamma_3 V_t^2 + \gamma_4 r_t^2 \right) r_t + \epsilon_t$$

Periodo e specifiche	γ_1 (s.e.)	γ_2 (s.e.)	γ_3 (s.e.)	γ_4 (s.e.)	R^2
A: Crisi Sub-Prime					
Volume	-0.208 (0.260)				0.029
Volatilità		0.013 (0.040)			0.028
Volume, Volatilità e r_t^2	0.104 (0.303)	0.009 (0.043)	-0.884 (0.648)	0.009 (0.010)	0.039
B: Post Crisi Subprime					
Volume	-0.418 (0.202)				0.046
Volatilità		-0.115 (0.101)			0.018
Volume, Volatilità, e r_t^2	0.405 (0.200)	0.017 (0.089)	-1.269 (0.351)	-0.002 (0.017)	0.046
C: Ripresa Economica					
Volume	0.161 (0.235)				0.019
Volatilità		-0.143 (0.127)			0.021
Volume, Volatilità e r_t^2	0.146 (0.296)	-0.299 (0.164)	0.808 (0.594)	0.035 (0.026)	0.032
D: Covid-19 e Guerre					
Volume	-0.647 (0.330)				0.099
Volatilità		-0.112 (0.058)			0.092
Volume, Volatilità e r_t^2	0.012 (0.481)	-0.013 (0.064)	-0.941 (1.132)	-0.015 (0.028)	0.110
E: Full Period					
Volume	-0.403 (0.160)				0.045
Volatilità		-0.060 (0.031)			0.030
Volume, Volatilità e r_t^2	0.169 (0.170)	-0.026 (0.029)	-0.883 (0.368)	0.000 (0.011)	0.045

Tabella 2.8: Volume e la Prima Autocorrelazione dei Rendimenti: Modello con Medie Mobili Triangolari

$$r_{t+1} = \alpha + \left(\gamma_{1(3m)}V_{t(3m)} + \gamma_{2(6m)}V_{t(6m)} + \gamma_{3(3y)}V_{t(3y)} + \sum_{i=1}^5 \beta_i D_i \right) r_t + \epsilon_t$$

Periodo e specifiche	$\gamma_{1(3m)}$ (s.e.)	$\gamma_{2(6m)}$ (s.e.)	$\gamma_{3(3y)}$ (s.e.)	R^2
A: Crisi Sub-Prime				
Volume (3 mesi)	-0.454 (0.275)			0.018
Volume (6 mesi)		-0.403 (0.249)		0.017
Volume (3 anni)			-0.398 (0.214)	0.019
B: Post Crisi Subprime				
Volume (3 mesi)	-0.531 (0.221)			0.028
Volume (6 mesi)		-0.474 (0.211)		0.028
Volume (3 anni)			-0.410 (0.192)	0.026
C: Ripresa Economica				
Volume (3 mesi)	-0.007 (0.268)			0.005
Volume (6 mesi)		-0.009 (0.236)		0.005
Volume (3 anni)			.0008 (0.202)	0.005
D: Covid-19 e Guerre				
Volume (3 mesi)	-0.849 (0.287)			0.083
Volume (6 mesi)		-0.727 (0.242)		0.091
Volume (3 anni)			-0.611 (0.177)	0.185
E: Full Period				
Volume (3 mesi)	-0.580 (0.150)			0.027
Volume (6 mesi)		-0.530 (0.137)		0.029
Volume (3 anni)			-0.489 (0.167)	0.036

il volume di trading e l'autocorrelazione del rendimento azionario. Quest'ultimo risultato conferma l'importanza del volume come fattore predittivo, con un impatto maggiore nel breve termine. Tuttavia, la bassa varianza spiegata, ossia il R^2 , suggerisce la necessità di considerare ulteriori variabili per una comprensione più completa dei rendimenti attesi.

2.6 Sviluppi Intertemporali: La Ricerca di Campbell

Per verificare la solidità di questi risultati, oltre a compararli con i risultati ottenuti per l'intervallo di ricerca di Campbell et al. la decisione è stata di proseguire la loro analisi dalla data di del loro documento, ossia dalla fine del 1988. In questo modo si ha una panoramica con data di inizio 1926 e data di fine 2023, quasi 100 anni di ricerca.

L'analisi dell'autocorrelazione dei rendimenti azionari nel corso dei vari periodi storici fornisce una panoramica su come diversi eventi economici e finanziari abbiano influenzato il comportamento del mercato.

Dai risultati, si osserva che i coefficienti di autocorrelazione (β) variano notevolmente tra i periodi. Nei primi anni di analisi (1926-1987), i coefficienti sono generalmente positivi, indicando che i rendimenti futuri tendono a essere influenzati positivamente dai rendimenti passati. Tuttavia, il valore degli R^2 è relativamente basso, suggerendo che la quantità di varianza nei rendimenti spiegata dai modelli è limitata.

Durante i periodi più recenti (1988-2023), si notano alcune variazioni significative. In particolare, i coefficienti di autocorrelazione diventano negativi in periodi di crisi finanziarie come la Crisi Finanziaria Globale (2003-2011) e il periodo del COVID-19 e oltre (2020-2023). Questo cambiamento indica che i rendimenti passati negativi tendono a predire rendimenti futuri negativi, riflettendo una maggiore volatilità e incertezza nei mercati finanziari durante questi periodi. Inoltre, gli R^2 aumentano leggermente nei periodi di crisi, suggerendo che l'autocorrelazione spiega una parte maggiore della varianza dei rendimenti in tempi di instabilità economica.

L'andamento dei coefficienti β può essere spiegato attraverso una lente economica. Nei periodi di crescita economica e stabilità, i mercati tendono a mostrare una maggiore continuità nei rendimenti, come indicato dai coefficienti positivi e dagli R^2 bassi. Tuttavia, durante i periodi di crisi, l'incertezza economica porta a movimenti più imprevedibili dei prezzi delle azioni, risultando in coefficienti negativi e R^2 leggermente più alti.

Questi risultati riflettono l'effetto dei fattori macroeconomici, come la politica monetaria, le crisi finanziarie globali e gli eventi straordinari come la pandemia di COVID-19, sui comportamenti dei mercati azionari. L'aumento della negatività dei coefficienti di autocorrelazione e l'incremento degli R^2 durante i periodi di crisi suggeriscono che i mercati sono più reattivi e volatili, influenzati pesantemente dalle condizioni economiche generali e dagli shock esterni.

Tabella 2.9: La Prima Autocorrelazione Dei Rendimenti Azionari: Un Campionamento Alternativo

$$r_{t+1} = \alpha + \beta r_t$$

$$r_{t+1} = \alpha + \left(\sum_{i=1}^5 \beta_i D_i \right) r_t$$

Periodo	β (<i>s.e.</i>)	R^2	Dummy R^2
Dal 1926-01-02 al 1962-06-29	0.039 (0.023)	0.002	0.005
Dal 1926-01-02 al 1939-12-30	0.015 (0.029)	0.000	0.004
Dal 1940-01-02 al 1949-12-31	0.112 (0.034)	0.012	0.018
Dal 1950-01-03 al 1962-06-29	0.130 (0.036)	0.017	0.037
Dal 1926-01-02 al 1987-09-30	0.073 (0.019)	0.005	0.008
Dal 1988-01-01 al 1992-12-31	0.010 (0.035)	0.000	0.002
Dal 1993-01-01 al 2002-12-31	0.001 (0.030)	0.000	0.000
Dal 2003-01-01 al 2011-12-31	-0.120 (0.037)	0.014	0.015
Dal 2012-01-01 al 2019-12-31	-0.017 (0.034)	0.000	0.002
Dal 2020-01-01 al 2023-12-31	-0.196 (0.079)	0.038	0.039
Dal 1988-01-01 al 2023-12-31	-0.079 (0.022)	0.006	0.007

Tabella 2.10: Volume, Volatilità, e la Prima Autocorrelazione dei Rendimenti: Un Campionamento Alternativo

$$r_{t+1} = \alpha + \left(\sum_{i=1}^5 \beta_i D_i + \gamma_1 V_t + \gamma_2 \sigma_t^2 + \gamma_3 V_t^2 \right) r_t + \epsilon_t$$

Periodo e specifiche	γ_1 (s.e.)	γ_2 (s.e.)	γ_3 (s.e.)	R^2
1926-1962				
Volume	0.053 (0.045)			0.006
Volatilità		-0.002 (0.041)		0.005
Volume e Volatilità	-0.114 (0.039)	0.064 (0.051)	0.009 (0.043)	0.007
1926-1939				
Volume	-0.038 (0.053)			0.004
Volatilità		0.027 (0.044)		0.004
Volume e Volatilità	-0.131 (0.047)	0.079 (0.058)	0.042 (0.046)	0.007
1940-1949				
Volume	-0.097 (0.061)			0.021
Volatilità		-0.173 (0.114)		0.021
Volume e Volatilità	-0.104 (0.062)	0.046 (0.064)	-0.130 (0.145)	0.022
1950-1962				
Volume	-0.174 (0.099)			0.042
Volatilità		-0.420 (0.199)		0.044
Volume e Volatilità	-0.152 (0.094)	0.216 (0.175)	-0.492 (0.244)	0.046
1926-1987				
Volume	-0.085 (0.043)			0.011
Volatilità		-0.029 (0.040)		0.009

Tabella 2.10: (continua)

Periodo e specifiche	γ_1 (s.e.)	γ_2 (s.e.)	γ_3 (s.e.)	R^2
Volume e Volatilità	-0.119 (0.039)	0.050 (0.049)	-0.014 (0.041)	0.011
1988-1992				
Volume	-0.134 (0.243)			0.013
Volatilità		0.124 (0.165)		0.013
Volume e Volatilità	-0.185 (0.226)	0.111 (0.196)	0.190 (0.382)	0.014
1993-2002				
Volume	-0.039 (0.189)			0.008
Volatilità		-0.020 (0.057)		0.008
Volume e Volatilità	0.073 (0.146)	-0.018 (0.059)	-0.530 (0.307)	0.013
2003-2011				
Volume	-0.256 (0.155)			0.025
Volatilità		-0.004 (0.035)		0.021
Volume e Volatilità	0.102 (0.190)	-0.002 (0.037)	-0.765 (0.251)	0.033
2012-2019				
Volume	0.230 (0.152)			0.013
Volatilità		-0.077 (0.109)		0.011
Volume e Volatilità	0.147 (0.208)	-0.153 (0.120)	0.324 (0.289)	0.018
2020-2023				
Volume	-0.592 (0.307)			0.090
Volatilità		-0.113 (0.057)		0.092
Volume e Volatilità	-0.157 (0.417)	-0.067 (0.053)	-0.238 (0.673)	0.097

Tabella 2.10: (continua)

Periodo e specifiche	γ_1 (s.e.)	γ_2 (s.e.)	γ_3 (s.e.)	R^2
1988-2023				
Volume	-0.291 (0.116)			0.019
Volatilità		-0.067 (0.029)		0.020
Volume e Volatilità	0.027 (0.083)	-0.043 (0.026)	-0.479 (0.170)	0.026

I risultati, inoltre, mostrano un'evoluzione interessante della relazione tra volume, volatilità e la prima autocorrelazione dei rendimenti nei vari periodi storici. Nel periodo considerato da Campbell (1926-1987), il volume ha un'influenza mista sui rendimenti attesi, con coefficienti di volume sia negativi che positivi. La volatilità tende ad avere un impatto negativo sui rendimenti, specialmente nel periodo post-1950.

Nel periodo recente (1988-2023), emergono differenze significative. I coefficienti del volume sono generalmente negativi, indicando che un aumento del volume è associato a una riduzione dei rendimenti attesi, particolarmente durante le crisi come la crisi finanziaria globale e la pandemia di COVID-19. Tuttavia, nel periodo post-crisi finanziaria (2012-2019), il volume ha un coefficiente positivo, suggerendo che in alcuni contesti un aumento del volume può essere associato a rendimenti più elevati.

La volatilità continua ad avere un impatto negativo sui rendimenti anche nel periodo recente. L'inclusione di entrambe le variabili di volume e volatilità nei modelli complessi mostra che le interazioni tra queste variabili possono influenzare significativamente i rendimenti attesi, come indicato dai coefficienti e dagli R^2 aumentati.

In sintesi, ci sono differenze significative tra i periodi storici, ma alcuni punti in comune rimangono. Entrambi i periodi mostrano che il volume e la volatilità sono importanti predittori dei rendimenti attesi, con il contesto economico e gli eventi specifici che influenzano l'entità e la direzione di questi effetti.

2.7 Modello di Markov Switching

Il modello di Markov Switching è uno strumento per analizzare serie temporali finanziarie che mostrano cambiamenti strutturali o regimi differenti nel tempo. Questo tipo di modello è utile per identificare e caratterizzare periodi con diverse dinamiche nei rendimenti, come fasi di alta e bassa volatilità, e per catturare eventi economici significativi che influenzano il comportamento del mercato¹¹.

¹¹Hamilton, J. D. (1989). "A new approach to the economic analysis of nonstationary time series and the business cycle." *Econometrica*, 57(2), 357-384.

2.7.1 Descrizione del Modello

Il modello di Markov Switching utilizzato in questa analisi specifica tre regimi ($k_{regimes} = 3$), con un'intercetta che varia tra i regimi e una varianza che cambia a seconda del regime¹².

2.7.2 Equazioni del Modello

Il modello di Markov Switching può essere descritto dalle seguenti equazioni:

Equazione dei Rendimenti

$$r_t = \mu_{S_t} + \epsilon_{t,S_t}$$

Dove:

- r_t è il rendimento logaritmico al tempo t .
- μ_{S_t} è l'intercetta (media) specifica del regime S_t .
- $\epsilon_{t,S_t} \sim N(0, \sigma_{S_t}^2)$ è l'errore con varianza specifica del regime S_t .

Processo di Markov

$$P(S_t = j \mid S_{t-1} = i) = p_{ij}$$

Dove:

- S_t è lo stato (regime) al tempo t .
- p_{ij} è la probabilità di transizione dallo stato i allo stato j .

Max Verosimiglianza

Il modello di Markov Switching viene stimato tramite la massimizzazione della log-verosimiglianza:

$$\mathcal{L}(\theta) = \sum_{t=1}^T \log \left(\sum_{j=1}^k P(S_t = j \mid S_{t-1}) f(r_t \mid \theta_j, S_t = j) \right)$$

Dove:

- $\mathcal{L}(\theta)$ è la funzione di log-verosimiglianza.
- $P(S_t = j \mid S_{t-1})$ è la probabilità di transizione.
- $f(r_t \mid \theta_j, S_t = j)$ è la densità condizionale dei rendimenti dati i parametri θ_j del regime j .

2.7.3 Risultati del Modello

I risultati del modello sono presentati nelle tabelle seguenti:

¹²Goldfeld, S. M., & Quandt, R. E. (1973). "A Markov model for switching regressions." *Journal of Econometrics*, 1(1), 3-15.

Tabella 2.11: Risultati del Modello di Markov Switching

	Regime 0	Regime 1	Regime 2
Intercetta	0.0944 (0.009)	-0.0050 (0.021)	-0.1985 (0.116)
Varianza (σ^2)	0.3130 (0.011)	1.3390 (0.054)	7.8609 (0.614)

Tabella 2.12: Parametri di Transizione del Modello di Markov Switching

Transizione	Parametro	Errore Standard
p[0→0]	0.9783	0.003
p[1→0]	0.0258	0.004
p[2→0]	0.0003	0.014
p[0→1]	0.0213	0.003
p[1→1]	0.9673	0.004
p[2→1]	0.0419	0.012

2.7.4 Regimi Identificati

Il modello ha identificato tre regimi distinti:

- **Regime 0:** Questo regime è caratterizzato da una **bassa varianza** ($\sigma^2 = 0.3130$) e una **intercetta positiva** ($\alpha = 0.0944$). Questo suggerisce un periodo di mercato relativamente stabile con rendimenti mediamente positivi.
- **Regime 1:** Questo regime mostra una **varianza intermedia** ($\sigma^2 = 1.3390$) e una **intercetta prossima a zero** ($\alpha = -0.0050$). Indica un periodo di mercato con volatilità moderata e rendimenti vicini allo zero, il che potrebbe essere indicativo di un mercato in equilibrio o in transizione.
- **Regime 2:** Questo regime è caratterizzato da una **alta varianza** ($\sigma^2 = 7.8609$) e una **intercetta negativa** ($\alpha = -0.1985$). Rappresenta un periodo di forte volatilità e rendimenti mediamente negativi, suggerendo condizioni di mercato avverse o di crisi.

2.7.5 Probabilità di Transizione

Le probabilità di transizione forniscono informazioni su quanto è probabile che il mercato passi da un regime all'altro:

- **p[0→0] = 0.9783:** Alta probabilità che il mercato resti nel regime 0, indicando una persistenza di condizioni di mercato stabili.
- **p[1→0] = 0.0258:** Bassa probabilità che il mercato transiti dal regime 1 al regime 0.
- **p[2→0] = 0.0003:** Quasi nulla probabilità di passaggio diretto dal regime 2 al regime 0.
- **p[0→1] = 0.0213:** Bassa probabilità di passaggio dal regime 0 al regime 1.

- $p[1 \rightarrow 1] = 0.9673$: Alta probabilità che il mercato resti nel regime 1, suggerendo che i periodi di volatilità moderata sono anche persistenti.
- $p[2 \rightarrow 1] = 0.0419$: Bassa probabilità di passaggio dal regime 2 al regime 1.

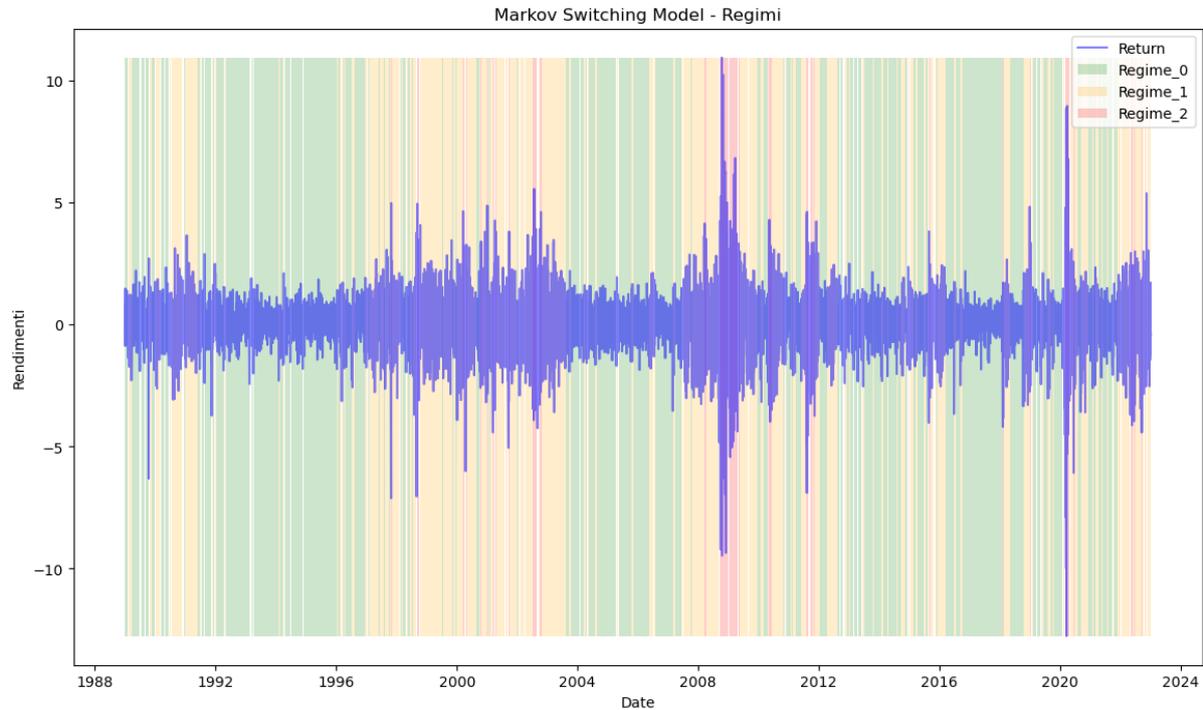


Figura 2.1: Panoramica sui regimi ed i rendimenti dello S&P 500

2.7.6 Impatto del Volume e Confronto con i Risultati Precedenti

Il volume di scambio gioca un ruolo cruciale nel contesto del modello di Markov Switching. Nei regimi identificati, il volume può influenzare i rendimenti in modo diverso, e i risultati ottenuti fino ad ora possono essere interpretati attraverso i regimi del modello di Markov.

- **Regime 0:** In questo regime caratterizzato da bassa volatilità e rendimenti positivi ($\alpha = 0.0944$), un aumento del volume potrebbe indicare un'accresciuta partecipazione al mercato. Tuttavia, dai risultati ottenuti precedentemente, abbiamo visto che l'impatto del volume sui rendimenti attesi tende spesso ad essere negativo. Questo potrebbe suggerire che, anche in un regime stabile, un aumento significativo del volume potrebbe essere dovuto a prese di profitto, portando a una pressione negativa sui rendimenti.
- **Regime 1:** Questo regime presenta volatilità intermedia e rendimenti vicini allo zero ($\alpha = -0.0050$). In tali condizioni, l'effetto del volume potrebbe essere neutro o ambiguo. I risultati precedenti hanno mostrato che l'impatto del volume può variare in base al periodo, ma generalmente non è fortemente positivo. In un regime di equilibrio o di transizione come questo, l'aumento del volume potrebbe non avere

un impatto definito sui rendimenti attesi, riflettendo l'incertezza e la mancanza di una direzione chiara del mercato.

- **Regime 2:** In questo regime di alta volatilità e rendimenti negativi ($\alpha = -0.1985$), un aumento del volume è spesso associato a vendite di panico o liquidazioni forzate. I risultati precedenti supportano questa interpretazione, mostrando che periodi di elevato volume sono frequentemente correlati a rendimenti negativi. In un contesto di crisi o di mercato avverso, un aumento del volume tende a peggiorare i rendimenti, confermando che in situazioni di alta incertezza e stress del mercato, il volume ha un impatto negativo significativo sui rendimenti attesi.

2.8 Trading Non Sincrono

Tutti i risultati fino ad ora ottenuti hanno utilizzato il rendimento di un indice ponderato per la capitalizzazione di mercato delle società che lo compongono. In generale, potrebbero risultare critici alcuni risultati in quanto si potrebbe obiettare che la correlazione seriale dei rendimenti dell'indice non sono tutti scambiati esattamente alla chiusura. In linea di principio il *trading non sincrono* si verifica quando i prezzi delle attività non vengono aggiornati simultaneamente sul mercato. Ciò può accadere per vari motivi, come la bassa liquidità di alcuni titoli, diverse ore di negoziazione tra mercati o ritardi nelle comunicazioni delle transazioni. A causa di questo fenomeno, i prezzi osservati possono riflettere informazioni obsolete rispetto ai prezzi di altri titoli, causando discrepanze temporanee nei dati di mercato¹³. Matematicamente, possiamo rappresentare i prezzi osservati P_t come una combinazione dei prezzi "veri" P_t^* e un ritardo d_t :

$$P_t = P_{t-d_t}^*$$

dove d_t rappresenta il ritardo temporale con cui il prezzo P_t viene aggiornato rispetto al prezzo "vero" P_t^* ¹⁴.

2.8.1 Impatto del Trading Non Sincrono sulla Correlazione Seriale dei Rendimenti

Il trading non sincrono può influenzare significativamente la correlazione seriale dei rendimenti, soprattutto in indici come l'S&P 500. La correlazione seriale dei rendimenti si riferisce alla relazione tra i rendimenti in periodi successivi. In presenza di trading non sincrono, i rendimenti calcolati possono riflettere informazioni ritardate, causando correlazioni spurie. Sia R_t il rendimento calcolato come:

$$R_t = \ln \left(\frac{P_t}{P_{t-1}} \right)$$

In presenza di trading non sincrono, i rendimenti osservati possono essere influenzati dai ritardi nei prezzi, introducendo correlazioni artificiali tra i rendimenti R_t e R_{t-1} . Questo può portare a una sovrastima o sottostima della correlazione seriale, distorcendo

¹³Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). "Trading Volume and Serial Correlation in Stock Returns." *Quarterly Journal of Economics*, 108(4), 905-939.

¹⁴Lo, A. W., & MacKinlay, A. C. (1990). "An econometric analysis of nonsynchronous trading." *Journal of Econometrics*, 45(1-2), 181-211.

le analisi statistiche e le inferenze economiche. Sebbene Lo e Mackinlay (1990) abbiano dimostrato che questo effetto è molto ridotto, a meno che i titoli non vengano scambiati per periodi di tempo inverosimilmente lunghi¹⁵. Un'ulteriore spunto di questa verifica è stata condotta pure da Campbell, Grossman e Wang nella ricerca di cui questo elaborato tende a replicare e supportare con i dati odierni¹⁶. Per rispondere a quest'obiezione, in linea con la ricerca di Campbell et al., l'analisi è stata svolta sul DOW JONES in merito alla correlazione seriale dei rendimenti dell'indice e della relazione che sussiste tra il rendimento atteso ed il volume, volatilità e le dummies. Sebbene questa serie ometta i dividendi, ciò ha solo un effetto minimo sul comportamento del DOW JONES. La scelta di analizzare il DOW JONES è fortemente razionale, poiché quest'ultimo è strutturalmente diverso dallo S&P 500. Il DJIA è composto da un numero relativamente piccolo di società (30) rispetto allo S&P 500 (500). Questo in prima analisi rende più semplice monitorare e analizzare individualmente ciascun componente dell'indice per effetti di trading non sincrone¹⁷. Inoltre, le società che compongono l'indice sono generalmente tra le più grandi e liquide del mercato. Ciò significa che i loro titoli vengono scambiati frequentemente e i prezzi sono aggiornati più rapidamente, riducendo il rischio di trading non sincrone. Di seguito, i risultati in linea con i modelli eseguito sino ad ora per l'autocorrelazione di primo ordine:

Tabella 2.13: Beta e R-Squared: Autocorrelazione di Primo Ordine del Dow Jones Industrial Average

$$r_{t+1} = \alpha + \left(\sum_{i=1}^5 \beta_i D_i + \gamma_1 V_t \right) r_t + \epsilon_t$$

Periodo e specifiche	β (s.e.)	R^2	Dummy R^2
Crisi Sub-Prime	-0.133 (0.052)	0.018	0.019
Post Crisi Subprime	-0.070 (0.052)	0.005	0.006
Ripresa Economica	-0.050 (0.037)	0.003	0.004
Covid-19 e Guerre	-0.205 (0.091)	0.042	0.044
Full Period	-0.130 (0.036)	0.017	0.017

¹⁵Lo, A. W., & MacKinlay, A. C. (1990). "An econometric analysis of nonsynchronous trading." *Journal of Econometrics*, 45(1-2), 181-211.

¹⁶Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). "Trading Volume and Serial Correlation in Stock Returns." *Quarterly Journal of Economics*, 108(4), 905-939.

¹⁷Lo, A. W., & MacKinlay, A. C. (1990). "An econometric analysis of nonsynchronous trading." *Journal of Econometrics*, 45(1-2), 181-211.

Commento sui Risultati del DOW JONES

I risultati mostrano che i coefficienti β dell'autocorrelazione di primo ordine sono generalmente negativi nei diversi periodi considerati. Questo suggerisce una tendenza generale verso un'inversione nei rendimenti dell'indice analizzato, coerente con i risultati ottenuti per l'S&P 500.

- **Crisi Sub-Prime:** Il β negativo e significativo indica una forte autocorrelazione negativa, suggerendo che i rendimenti tendono a invertire la direzione in questo periodo.
- **Post Crisi Subprime:** Il β è ancora negativo ma meno pronunciato rispetto al periodo della Crisi Sub-Prime, indicando una minore autocorrelazione negativa.
- **Ripresa Economica:** Il β diventa meno negativo e quasi nullo, suggerendo una diminuzione dell'autocorrelazione negativa durante la ripresa economica.
- **Covid-19 e Guerre:** Il β torna ad essere fortemente negativo, riflettendo una maggiore autocorrelazione negativa durante periodi di alta incertezza e volatilità.
- **Full Period:** Il β negativo per l'intero periodo conferma la tendenza generale all'autocorrelazione negativa nei rendimenti dell'indice.

Nel complesso, i risultati sono coerenti con l'analisi dell'S&P 500, mostrando una significativa autocorrelazione negativa nei rendimenti durante periodi di crisi e alta volatilità. I risultati presentati nella tabella 16 mostrano i coefficienti del volume, della volatilità e del volume quadratico ($\gamma_1, \gamma_2, \gamma_3$) insieme agli R^2 corrispondenti per le diverse specifiche del modello, applicato al Dow Jones Industrial Average (DJIA). Confrontando questi risultati con quelli ottenuti per l'S&P 500, emergono alcune somiglianze e differenze interessanti:

- **Coefficiente del Volume (γ_1):** Per entrambi gli indici, i coefficienti del volume tendono ad essere negativi durante i periodi di crisi (Crisi Sub-Prime, Post Crisi Subprime, Covid-19 e Guerre).
- **Coefficiente della Volatilità (γ_2):** La volatilità ha un impatto negativo sui rendimenti per entrambi gli indici in quasi tutti i periodi. Questo risultato è coerente con l'idea che una maggiore incertezza di mercato tende a ridurre i rendimenti attesi.
- **Coefficiente del Volume Quadratico (γ_3):** L'inclusione del termine quadratico del volume migliora leggermente il potere esplicativo del modello per entrambi gli indici, ma con variazioni nei segni e nelle magnitudini dei coefficienti.
- **R-Squared (R^2):** I valori di R^2 indicano che il potere esplicativo dei modelli è relativamente basso per entrambi gli indici, ma il modello applicato al DJIA sembra avere un R^2 leggermente superiore in alcuni periodi, come nel periodo di Covid-19 e Guerre.

In sintesi, i risultati per il DJIA sono in gran parte coerenti con quelli dell'S&P 500, specialmente per quanto riguarda l'impatto negativo del volume e della volatilità durante i periodi di crisi.

Tabella 2.14: Volume, Volatilità, e la Prima Autocorrelazione dei Rendimenti: Dow Jones Industrial Average

$$r_{t+1} = \alpha + \left(\sum_{i=1}^5 \beta_i D_i + \gamma_1 V_t + \gamma_2 \sigma_t^2 + \gamma_3 V_t^2 \right) r_t + \epsilon_t$$

Periodo e specifiche	γ_1 (s.e.)	γ_2 (s.e.)	γ_3 (s.e.)	R^2
Crisi Sub-Prime				
Volume	-0.075 (0.211)			0.025
Volatilità		-0.013 (0.043)		0.025
Volume e Volatilità	0.067 (0.230)	0.010 (0.047)	-0.338 (0.533)	0.027
Post Crisi Subprime				
Volume	-0.395 (0.204)			0.033
Volatilità		-0.150 (0.144)		0.023
Volume e Volatilità	0.358 (0.205)	0.010 (0.118)	-1.079 (0.318)	0.060
Ripresa Economica				
Volume	-0.078 (0.127)			0.017
Volatilità		-0.145 (0.124)		0.020
Volume e Volatilità	-0.042 (0.172)	-0.156 (0.124)	0.163 (0.345)	0.020
Covid-19 e Guerre				
Volume	-0.504 (0.301)			0.106
Volatilità		-0.091 (0.053)		0.105
Volume e Volatilità	-0.062 (0.549)	-0.040 (0.066)	-0.299 (0.756)	0.111
Full Period				
Volume	-0.311 (0.149)			0.036
Volatilità		-0.066 (0.034)		0.036
Volume e Volatilità	0.106 (0.144)	-0.031 (0.031)	-0.471 (0.276)	0.044

2.8.2 Modello alternativo per il Trading non Sincrono: Analisi dei rendimenti individuali

Un altro modo per rispondere al problema del trading non sincrono è quello di utilizzare i dati sui rendimenti dei singoli titoli. Tuttavia, le negoziazioni non sincrone hanno solo un effetto trascurabile sulle autocorrelazioni dei rendimenti dei singoli titoli¹⁸. Anche quando si utilizzano i rendimenti dei singoli titoli il volume aggregato è probabilmente una metrica migliore rispetto al volume individuale, in quanto la pressione idiosincratca di acquisto o di vendita non crea un rischio sistematico per i market maker¹⁹.

Per analizzare l'effetto dei volumi e delle variabili temporali sui rendimenti azionari, sono stati considerati tre modelli di regressione: regressione sull'indice equal-weighted, regressione pooled e regressioni individuali. Questi modelli sono stati formulati come segue:

Regressione sull'Indice Equal-Weighted

La regressione per l'indice equal-weighted è definita da:

$$r_{EW,t+1} = \alpha_{EW} + \left(\sum_{i=1}^5 \beta_{i,EW} D_{i,t} + \gamma_{EW} V_t \right) r_{EW,t} + \epsilon_t \quad (2.1)$$

Dove:

- $r_{EW,t+1}$ è il rendimento dell'indice equal-weighted al tempo $t + 1$.
- α_{EW} è l'intercetta della regressione.
- $\beta_{i,EW}$ sono i coefficienti per le variabili dummy settimanali $D_{i,t}$ moltiplicate per il rendimento al tempo t .
- γ_{EW} è il coefficiente del turnover del volume V_t moltiplicato per il rendimento $r_{EW,t}$.
- ϵ_t è il termine di errore, distribuito come $\mathcal{N}(0, \sigma_\epsilon^2)$.

Questa regressione mira a captare l'effetto combinato dei volumi e delle variazioni settimanali sui rendimenti dell'indice, considerando che le variabili temporali possono influenzare i pattern di trading.

Regressione Pooled

Nella regressione pooled, i rendimenti individuali delle azioni sono combinati in un singolo modello:

$$r_{j,t+1} = \alpha_p + \left(\sum_{i=1}^5 \beta_{i,p} D_{i,t} + \gamma_p V_t \right) r_{j,t} + \epsilon_{j,t} \quad (2.2)$$

Dove:

- $r_{j,t+1}$ è il rendimento dell'azione j al tempo $t + 1$.

¹⁸Lo, A. W., & MacKinlay, A. C. (1990). "An econometric analysis of nonsynchronous trading." *Journal of Econometrics*, 45(1-2), 181-211.

¹⁹Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). "Trading Volume and Serial Correlation in Stock Returns." *Quarterly Journal of Economics*, 108(4), 905-939.

- α_p è l'intercetta della regressione pooled.
- $\beta_{i,p}$ sono i coefficienti per le variabili dummy settimanali $D_{i,t}$ moltiplicate per il rendimento al tempo t .
- γ_p è il coefficiente del turnover del volume V_t moltiplicato per il rendimento $r_{j,t}$.
- $\epsilon_{j,t}$ è il termine di errore, distribuito come $\mathcal{N}(0, \sigma_\epsilon^2)$.

Questa regressione combina i dati di tutte le azioni per stimare un effetto medio del volume e delle variabili temporali sui rendimenti, considerando che gli effetti osservati siano comuni a tutte le azioni.

Regressioni Individuali

Le regressioni individuali sono stimate separatamente per ogni azione:

$$r_{j,t+1} = \alpha_j + \left(\sum_{i=1}^5 \beta_{i,j} D_{i,t} + \gamma_j V_t \right) r_{j,t} + \epsilon_{j,t} \quad (2.3)$$

Dove:

- α_j è l'intercetta della regressione per l'azione j .
- $\beta_{i,j}$ sono i coefficienti per le variabili dummy settimanali $D_{i,t}$ moltiplicate per il rendimento al tempo t per l'azione j .
- γ_j è il coefficiente del turnover del volume V_t moltiplicato per il rendimento $r_{j,t}$ per l'azione j .
- $\epsilon_{j,t}$ è il termine di errore specifico per l'azione j , distribuito come $\mathcal{N}(0, \sigma_\epsilon^2)$.

Questo approccio permette di catturare le specificità di ciascuna azione, permettendo di identificare effetti idiosincratichi del volume e delle variabili settimanali sui rendimenti²⁰. La scelta delle società sulle quali analizzare gli effetti dei volumi sui rendimenti azionari è stata principalmente basata sulle dimensioni e sulla liquidità. A tal fine sono state riprese, per comparare i risultati ottenuti, le 30 maggiori società dell'azionariato statunitense nel periodo 2004-2023²¹. L'indice equamente ponderato di questi pattern di titoli ha una correlazione con il DOW JONES pari a 0,98. Non sorprende quindi che l'effetto del volume sull'autocorrelazione dell'indice sia simile a quello riportato nella Tabella 2.14. La terza colonna della tabella 2.15 mostra l'effetto del volume sulla correlazione di ciascun rendimento azionario con il proprio primo ritardo, dove i singoli rendimenti sono sovrapposti in un'unica regressione in pool.

²⁰Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). "Trading Volume and Serial Correlation in Stock Returns." *Quarterly Journal of Economics*, 108(4), 905-939.

²¹Le Società scelte sono le seguenti: Apple Inc. - AAPL, Microsoft Corporation - MSFT, Amazon.com Inc. - AMZN, Verizon Communications Inc - VZ, Berkshire Hathaway Inc. - BRK/A, BRK/B, Johnson & Johnson - JNJ, JPMorgan Chase & Co. - JPM, Exxon Mobil Corporation - XOM, Procter & Gamble Co. - PG, Walmart Inc. - WMT, UnitedHealth Group - UNH, General Electric Company - MA, American Express Company - AXP, Intel Corporation - INTC, AT&T Inc. - T, Coca-Cola Company - KO, PepsiCo, Inc. - PEP, Merck & Co., Inc. - MRK, Walt Disney Company - DIS, Cisco Systems, Inc. - CSCO, Chevron Corporation - CVX, Pfizer Inc. - PFE, Bank of America Corporation - BAC, IBM (International Business Machines Corporation) - IBM, McDonald's Corporation - MCD, 3M Company - MMM, Bristol-Myers Squibb Company - BMY, Abbott Laboratories - ABT, General Electric Company - GE, Oracle Corporation - ORCL.

Tabella 2.15: Effetto del Volume e delle Variabili Temporali sui Rendimenti Azionari

Equal-Weighted Index Regression:

$$r_{EW,t+1} = \alpha_{EW} + \left(\sum_{i=1}^5 \beta_{i,EW} D_{i,t} + \gamma_{EW} V_t \right) r_{EW,t} + \epsilon_t$$

Pooled Regression:

$$r_{j,t+1} = \alpha_p + \left(\sum_{i=1}^5 \beta_{i,p} D_{i,t} + \gamma_p V_t \right) r_{j,t} + \epsilon_{j,t}, \quad j = 1, \dots, 30$$

Individual Stock Regressions:

$$r_{j,t+1} = \alpha_j + \left(\sum_{i=1}^5 \beta_{i,j} D_{i,t} + \gamma_j V_t \right) r_{j,t} + \epsilon_{j,t}, \quad j = 1, \dots, 30$$

Periodo e specifiche	γ_{EW} (s.e.)	R_{EW}^2	γ_P (s.e.)	R_P^2	γ_j ($\gamma < 0$)	\bar{t} ($t < 0$)
Crisi Sub-Prime	-0.025 (0.203)	0.030	-0.013 (0.028)	0.028	-0.012 (10)	-0.163 (1)
Post Crisi Sub-Prime	-0.443 (0.206)	0.035	-0.145 (0.017)	0.018	-0.150 (28)	-0.996 (4)
Ripresa Economica	0.159 (0.134)	0.015	0.023 (0.017)	0.009	0.022 (11)	0.209 (3)
Covid-19 e Guerre	-0.594 (0.337)	0.195	-0.117 (0.032)	0.106	-0.117 (24)	-0.620 (4)
Full Period	-0.054 (0.063)	0.034	-0.062 (0.015)	0.020	-0.062 (25)	-0.673 (5)

I risultati delle regressioni mostrano variazioni significative nell'effetto del volume sui rendimenti azionari attraverso i diversi periodi considerati.

- **Crisi Sub-Prime:** Il coefficiente γ_{EW} è negativo ma non molto pronunciato, indicando un effetto moderato del volume sui rendimenti. Tuttavia, i bassi valori di R^2 suggeriscono una capacità limitata del modello di spiegare la variabilità dei rendimenti in questo periodo.

- **Post Crisi Sub-Prime:** Il coefficiente γ_{EW} è significativamente negativo, indicando un impatto maggiore del volume sui rendimenti rispetto al periodo della crisi Sub-Prime. Il R^2 rimane relativamente basso, suggerendo che altri fattori potrebbero influenzare i rendimenti.

- **Ripresa Economica:** Durante questo periodo, il coefficiente γ_{EW} diventa positivo, indicando che un aumento del volume è associato a rendimenti più elevati. Tuttavia, l'effetto sembra essere debole, come indicato dal basso valore di R^2 .

- **Covid-19 e Guerre:** In questo periodo, il coefficiente γ_{EW} è fortemente negativo, indicando un significativo impatto negativo del volume sui rendimenti. Questo è coerente con l'elevata incertezza e volatilità che hanno caratterizzato i mercati durante la pandemia e altre turbolenze globali.

- **Full Period:** Il coefficiente γ_{EW} aggregato per l'intero periodo è negativo, suggeren-

do che l'aumento del volume è generalmente associato a rendimenti inferiori. Tuttavia, i valori di R^2 suggeriscono che il volume spiega solo una piccola parte della variabilità dei rendimenti. Questi risultati sono in linea con la letteratura esistente, che evidenzia che l'aumento del volume è spesso associato a un'elevata volatilità e a rendimenti attesi inferiori, specialmente in periodi di crisi o incertezza economica.

Capitolo 3

Volume e Correlazione: Un Modello Teorico

In questo capitolo, verrà presentato e discusso un modello teorico che esplora la relazione tra il volume di trading e la correlazione dei rendimenti nei mercati finanziari. Questo modello si basa sul lavoro pionieristico di Campbell, Grossman e Wang (1993), il cui contributo ha rappresentato un punto di riferimento nella letteratura finanziaria. La scelta di riprendere e adattare il modello di Campbell è motivata dalla sua capacità di spiegare dinamiche complesse nei mercati finanziari attraverso un quadro teorico rigoroso e quantitativo.

Il modello di Campbell si distingue per l'approccio innovativo nell'analizzare come le variazioni del volume di scambio possano influenzare e predire la serialità dei rendimenti delle attività finanziarie. In particolare, il modello incorpora elementi chiave come l'avversione al rischio degli investitori, la componente stocastica dei dividendi e il comportamento del prezzo delle attività rischiose. Questi elementi vengono utilizzati per costruire un quadro economico che permette di comprendere meglio le forze che guidano la formazione dei prezzi e la loro volatilità.

Nel corso di questo capitolo, verranno esaminati i fondamenti teorici del modello, seguiti dalla sua implementazione e dall'analisi dei risultati. Verranno inoltre discusse le ipotesi sottostanti e le possibili estensioni del modello, con l'obiettivo di fornire una comprensione approfondita delle dinamiche che esso descrive.

L'adozione del modello di Campbell non è solo un omaggio alla sua rilevanza storica, ma anche un tentativo di applicare strumenti teorici consolidati a contesti contemporanei, al fine di verificare la loro validità e utilità nel descrivere fenomeni finanziari moderni.

3.1 Struttura dell'Economia

Per comprendere quanto meglio il modello è necessario dare una spiegazione chiara dei vari passaggi e assunzioni su cui poggia quest'ultimo poiché sarà di fondamentale importanza per comprenderne i risultati successivamente esposti. In questo caso, il modello teorico considera un'economia in cui esiste un *asset privo di rischio* con un tasso di rendimento garantito pari a $R = 1 + r$, dove $r > 0$. Questo asset è disponibile in offerta elastica, consentendo agli investitori di ottenere un ritorno costante e sicuro.

Si assume che ci sia una *offerta fissa di azioni* per capita, normalizzata a 1. Ogni azione paga un dividendo D_t nel periodo t , che è composto da due componenti:

$$D_t = \bar{D} + \tilde{D}_t$$

- $\bar{D} > 0$ è il dividendo medio costante.
- \tilde{D}_t è la componente stocastica, che ha media zero e riflette le fluttuazioni casuali attorno al valore medio.

3.1.1 Processo Stocastico dei Dividendi

La componente stocastica del dividendo \tilde{D}_t segue un processo autoregressivo di ordine 1 (AR(1)):

$$\tilde{D}_t = \alpha_D \tilde{D}_{t-1} + u_{D,t} \quad (1)$$

dove:

- α_D è il coefficiente di autoregressione, con $0 \leq \alpha_D \leq 1$. Questo parametro misura l'influenza del valore passato del dividendo stocastico sul valore corrente.
 - Se $\alpha_D = 0$, \tilde{D}_t è un processo di puro rumore bianco, senza memoria dei valori passati.
 - Se $\alpha_D = 1$, \tilde{D}_t segue un processo *random walk*, dove ogni valore futuro è completamente determinato dal valore precedente più uno shock stocastico.
- $u_{D,t}$ è il termine di innovazione, distribuito normalmente con media zero e varianza σ_u^2 , ovvero $u_{D,t} \sim \mathcal{N}(0, \sigma_u^2)$.

Questo processo indica che il valore corrente di \tilde{D}_t dipende in parte dal valore passato, con un elemento di casualità introdotto dal termine di innovazione $u_{D,t}$.

3.1.2 Tipi di Investitori

Nell'economia descritta, si assumono due tipi di investitori:

- *Investitori di tipo A*: Questi investitori hanno un'avversione al rischio costante, descritta dal parametro a .
- *Investitori di tipo B*: Anch'essi hanno un'avversione al rischio, descritta dal parametro b_t , che però può variare nel tempo.

La frazione di investitori di tipo A è indicata con ω . La distinzione tra questi due tipi di investitori è cruciale per comprendere la dinamica delle decisioni di trading e l'evoluzione dei prezzi delle azioni in risposta alle fluttuazioni dei dividendi e alle variazioni dell'avversione al rischio.

3.2 Utilità Attesa, Prezzo e Segnale

Nel modello considerato, gli investitori massimizzano la loro utilità attesa, rappresentata da una funzione di utilità esponenziale negativa:

$$\max_X \mathbb{E}_t [-\exp(-\Psi W_{t+1})], \quad \Psi = ab_t \quad (2)$$

Questa funzione di utilità segue il modello **CARA (Constant Absolute Risk Aversion)**¹, che implica un'avversione al rischio costante in termini assoluti. Il modello CARA è descritto dalla funzione:

$$U(W) = -\exp(-\lambda W)$$

dove λ è il coefficiente di avversione al rischio. La ricchezza futura W_{t+1} è determinata dal vincolo di bilancio:

$$W_{t+1} = W_t R + X_t (P_{t+1} + D_{t+1} - RP_t) \quad (3)$$

dove:

- W_t è la ricchezza attuale al tempo t .
- R è il tasso di rendimento dell'asset privo di rischio.
- X_t è il numero di unità dell'asset rischioso detenute al tempo t .
- P_{t+1} è il prezzo dell'asset rischioso al tempo $t + 1$.
- D_{t+1} è il dividendo pagato dall'asset rischioso al tempo $t + 1$.
- RP_t rappresenta il costo di opportunità di detenere l'asset rischioso rispetto all'asset privo di rischio.

3.2.1 Massimizzazione dell'Utilità Attesa

1. Espressione del Vincolo di Bilancio:

$$W_{t+1} = W_t R + X_t (P_{t+1} + D_{t+1} - RP_t)$$

2. Funzione di Utilità:

$$U(W_{t+1}) = -\exp(-\Psi [W_t R + X_t (P_{t+1} + D_{t+1} - RP_t)])$$

3. Calcolo dell'Utilità Attesa:

$$\mathbb{E}_t [U(W_{t+1})] = \mathbb{E}_t [-\exp(-\Psi [W_t R + X_t (P_{t+1} + D_{t+1} - RP_t)])]$$

4. Derivazione di Primo Ordine:

$$\frac{\partial}{\partial X_t} \mathbb{E}_t [U(W_{t+1})] = \mathbb{E}_t [(P_{t+1} + D_{t+1} - RP_t) \exp(-\Psi [W_t R + X_t (P_{t+1} + D_{t+1} - RP_t)])] = 0$$

¹, K. J. (1971). "Essays in the Theory of Risk-Bearing". Chicago: Markham Publishing Company.

5. Soluzione della Massimizzazione:

Assumendo che la distribuzione dei rendimenti sia normale, possiamo risolvere per X_t :

$$X_t^* = \frac{\mathbb{E}_t [P_{t+1} + D_{t+1} - RP_t]}{\Psi \text{Var}(P_{t+1} + D_{t+1} - RP_t)} \quad (4)$$

3.2.2 Set Informativo e Segnale

Il set informativo \mathcal{I}_t disponibile agli investitori contiene:

- Il prezzo dell'azione P_t .
- Il dividendo D_t .
- Un segnale S_t , ricevuto da tutti gli investitori al tempo t , che fornisce informazioni riguardo al futuro shock ai dividendi $u_{D,t+1}$.

L'equazione che descrive il futuro shock ai dividendi è:

$$u_{D,t+1} = S_t + \epsilon_{D,t+1} \quad (5)$$

dove:

- S_t è un segnale comune a tutti gli investitori distribuita normalmente (i.i.d), ovvero $S_t \sim \mathcal{N}(0, \sigma_s^2)$, che rappresenta un'informazione anticipata sul futuro dividendo.
- $\epsilon_{D,t+1}$ è un errore idiosincratice distribuito normalmente con media zero e varianza σ_ϵ^2 , ovvero $\epsilon_{D,t+1} \sim \mathcal{N}(0, \sigma_\epsilon^2)$.

3.2.3 Prezzo di Equilibrio dell'Asset Rischioso

Il prezzo dell'asset rischioso F_t è determinato come il valore attuale atteso dei flussi di cassa futuri:

$$F_t = \mathbb{E}_t \left[\sum_{s=0}^{\infty} \frac{1}{R^s} D_{t+s} \middle| \tilde{D}_t, S_t \right] = \frac{R\bar{D}}{r} + \frac{R}{R - \alpha_D} \tilde{D}_t + \frac{1}{R - \alpha_D} S_t \quad (6)$$

Descrizione:

- **Valore Fondamentale:** F_t rappresenta il valore fondamentale dell'asset rischioso, calcolato come la somma scontata dei dividendi futuri attesi.
- **Componente Stocastica:** \tilde{D}_t e S_t influenzano il prezzo, rappresentando rispettivamente la componente stocastica del dividendo e il segnale informativo sul futuro.

3.2.4 Varianza dell'Innovazione del Prezzo

La varianza dell'innovazione del prezzo $\sigma_{F_t}^2$ è determinata da due componenti principali: la varianza dello shock idiosincratice $\epsilon_{D,t+1}$ e la varianza del segnale comune S_t .

Definizione del Prezzo di Equilibrio

Il prezzo dell'asset rischioso F_t è determinato come il valore attuale atteso dei dividendi futuri:

$$F_t = \mathbb{E}_t \left[\sum_{s=0}^{\infty} \frac{1}{R^s} D_{t+s} \mid \tilde{D}_t, S_t \right]$$

In questo caso, possiamo esprimere il dividendo D_{t+s} come:

$$D_{t+s} = \bar{D} + \tilde{D}_{t+s} \quad (7)$$

dove \tilde{D}_{t+s} segue il processo stocastico autoregressivo:

$$\tilde{D}_{t+s} = \alpha_D^s \tilde{D}_t + \sum_{j=0}^{s-1} \alpha_D^j u_{D,t+s-j} \quad (8)$$

Innovazione nel Prezzo dell'Asset

Il prezzo dell'asset F_t può essere riscritto come una somma di componenti legate al dividendo medio, alla componente stocastica del dividendo e al segnale:

$$F_t = \frac{R\bar{D}}{r} + \frac{R}{R - \alpha_D} \tilde{D}_t + \frac{1}{R - \alpha_D} S_t$$

Il termine $\frac{R\bar{D}}{r}$ è costante, quindi non contribuisce alla varianza dell'innovazione. Le due componenti rilevanti sono \tilde{D}_t e S_t .

Varianza dell'Innovazione del Prezzo

La varianza dell'innovazione del prezzo $\sigma_{F_t}^2$, dato che le variabili seguono un processo stocastico e vengono definite i.i.d., è data dalla somma delle varianze delle due componenti stocastiche:

$$\sigma_{F_t}^2 = \text{Var} \left(\frac{R}{R - \alpha_D} \tilde{D}_t \right) + \text{Var} \left(\frac{1}{R - \alpha_D} S_t \right)$$

Varianza della Componente \tilde{D}_t : La componente \tilde{D}_t segue un processo autoregressivo di primo ordine (AR(1)):

$$\tilde{D}_t = \alpha_D \tilde{D}_{t-1} + u_{D,t}$$

Con $u_{D,t} = S_t + \epsilon_{D,t+1}$. La varianza di \tilde{D}_t è data da:

$$\text{Var}(\tilde{D}_t) = \frac{\sigma_u^2}{1 - \alpha_D^2}$$

Dove σ_u^2 è la varianza di $u_{D,t}$. Dato che $u_{D,t} = S_t + \epsilon_{D,t+1}$, la varianza di $u_{D,t}$ è la somma della varianza di S_t e $\epsilon_{D,t+1}$:

$$\sigma_u^2 = \sigma_S^2 + \sigma_\epsilon^2$$

Quindi, la varianza della componente \tilde{D}_t nel prezzo F_t è:

$$\text{Var} \left(\frac{R}{R - \alpha_D} \tilde{D}_t \right) = \frac{R^2 \sigma_\epsilon^2}{(R - \alpha_D)^2}$$

Varianza della Componente S_t : La varianza della componente legata al segnale S_t nel prezzo F_t è semplicemente:

$$\text{Var} \left(\frac{1}{R - \alpha_D} S_t \right) = \frac{\sigma_S^2}{(R - \alpha_D)^2}$$

Somma delle Varianze

Sommando le due componenti, dato che sono indipendenti e identicamente distribuite, otteniamo la varianza totale dell'innovazione del prezzo F_t :

$$\sigma_{F_t}^2 = \frac{R^2 \sigma_\epsilon^2}{(R - \alpha_D)^2} + \frac{\sigma_S^2}{(R - \alpha_D)^2} \quad (9)$$

Nel caso in cui gli investitori siano neutrali al rischio, il prezzo ex-dividendo di equilibrio dell'azione è semplicemente dato da $F_t - D_t$:

$$P_t^{\text{ex-dividend}} = F_t - D_t$$

Tuttavia, quando gli investitori sono avversi al rischio, il prezzo di equilibrio dipende dall'avversione al rischio del mercato. In questo scenario, la varianza dell'innovazione del prezzo $\sigma_{F_t}^2$ gioca un ruolo cruciale nel determinare il prezzo di equilibrio, poiché gli investitori richiedono un premio per il rischio per detenere l'asset rischioso.

3.3 Avversione al Rischio e Determinazione del Prezzo

3.3.1 Definizione della Variabile Z_t

Definiamo la variabile Z_t , che può essere interpretata come l'avversione al rischio dell'investitore marginale nel mercato:

$$Z_t \equiv \frac{ab_t}{(1 - \omega)a + \omega b_t} \quad (10)$$

Dove:

- a è il parametro di avversione al rischio degli investitori di tipo A, costante nel tempo;
- b_t è il parametro di avversione al rischio degli investitori di tipo B, che può variare nel tempo;
- ω è la frazione del mercato rappresentata dagli investitori di tipo A.

3.3.2 Processo AR(1) per \tilde{Z}_t

La variabile Z_t può essere decomposta in due componenti:

$$Z_t = \bar{Z} + \tilde{Z}_t$$

Dove:

- \bar{Z} rappresenta una componente costante, ossia l'avversione al rischio aggregata media;
- \tilde{Z}_t è la componente stocastica che segue un processo autoregressivo di primo ordine (AR(1)):

$$\tilde{Z}_t = \alpha_Z \tilde{Z}_{t-1} + u_{Z,t}, \quad 0 \leq \alpha_Z \leq 1 \quad (11)$$

Dove:

- α_Z è il coefficiente di autoregressione. Valori di α_Z vicini a 1 indicano una persistenza elevata, mentre valori vicini a 0 indicano una scarsa persistenza;
- $u_{Z,t}$ è un termine di errore stocastico indipendente e identicamente distribuito (i.i.d.) con distribuzione normale, $u_{Z,t} \sim \mathcal{N}(0, \sigma_Z^2)$.

3.3.3 Assunzioni sul Processo Z_t

Si assume che $u_{Z,t}$ sia indipendente dagli altri shock e distribuito normalmente con varianza σ_Z^2 . Questa assunzione permette alla componente stocastica \tilde{Z}_t e quindi a b_t di essere negativi in alcuni casi. Tuttavia, la negatività di b_t potrebbe essere evitata sostituendo l'assunzione di utilità esponenziale con una funzione obiettivo che massimizza l'utilità attesa ponderata dalla varianza. Inoltre, si assume che la varianza σ_Z^2 sia limitata da una soglia superiore definita come σ_Z^{*2} :

$$\sigma_Z^2 \leq \sigma_Z^{*2} = \frac{(R - \alpha_Z)^2}{4\sigma_{F_t}^2}$$

Questo parametro σ_Z^{*2} rappresenta un limite superiore per la varianza del processo \tilde{Z}_t , oltre il quale il modello potrebbe non essere sostenibile o produrre risultati non realistici.

3.3.4 Prezzo di Equilibrio dell'Asset

Il primo teorema definito da Campbell, Grossman e Wang afferma che, data l'economia definita, esiste un prezzo di equilibrio per l'azione che è dato da:

$$P_t = F_t - D_t + (p_0 + p_Z Z_t) \quad (12)$$

Dove:

- $p_Z = -\frac{(R - \alpha_Z)}{2\sigma_Z^2} \left[1 - \sqrt{1 - \left(\frac{\sigma_Z^2}{\sigma_Z^{*2}} \right)} \right]$;
- $p_0 = \frac{(1 - \alpha_Z)p_Z \bar{Z}}{r} < 0$.

Questa formula mostra che il prezzo di equilibrio P_t dipende non solo dal valore fondamentale F_t e dai dividendi D_t , ma anche dall'avversione al rischio aggregata del mercato Z_t , attraverso i parametri p_0 e p_Z . Il termine p_0 è negativo, il che implica che l'avversione al rischio riduce il prezzo di equilibrio rispetto a un contesto di investitori neutrali al rischio.

3.4 Rendimento in Eccesso e Volume

3.4.1 Rendimento in Eccesso

Il rendimento in eccesso Q_{t+1} per un'azione tra i tempi t e $t + 1$ è definito come:

$$Q_{t+1} = P_{t+1} + D_{t+1} - RP_t \quad (13)$$

Dove:

- P_t è il prezzo dell'azione al tempo t ;
- P_{t+1} è il prezzo dell'azione al tempo $t + 1$;
- D_{t+1} è il dividendo pagato dall'azione nel periodo $[t, t + 1]$;
- R è il rendimento privo di rischio applicato al prezzo P_t .

Questa equazione riflette il rendimento totale dell'azione, incluso il guadagno in conto capitale (dovuto alla variazione del prezzo) e il rendimento da dividendo, al netto del rendimento privo di rischio.

Varianza del Rendimento in Eccesso

La varianza del rendimento in eccesso σ_Q^2 è determinata dalle principali componenti stocastiche che influenzano il modello. In particolare, essa è la somma della varianza del valore fondamentale F_t e della varianza legata all'avversione al rischio Z_t :

$$\sigma_Q^2 = \sigma_{F_t}^2 + p_Z^2 \sigma_Z^2 \quad (14)$$

Dove:

- $\sigma_{F_t}^2$ è la varianza del valore fondamentale F_t , che dipende dai dividendi futuri e dal segnale S_t .
- σ_Z^2 è la varianza del processo stocastico Z_t , che rappresenta l'avversione al rischio del mercato.
- p_Z è un coefficiente che riflette l'impatto dell'avversione al rischio sul rendimento in eccesso.

Questa espressione mostra come la varianza del rendimento in eccesso sia composta dalla somma delle varianze delle principali componenti del modello.

Espressione di Q_{t+1} in Termini di Innovazioni

Infine, possiamo riscrivere Q_{t+1} evidenziando le innovazioni inattese:

$$Q_{t+1} = e_t + \kappa[e_{t+1} - \mathbb{E}_t[e_{t+1}]] + (F_{t+1} - \mathbb{E}_t[F_{t+1}]) \quad (15)$$

Dove:

- $e_t = \mathbb{E}_t[Q_{t+1}|\mathcal{I}_t] = \sigma_Q^2 Z_t$ rappresenta il rendimento in eccesso atteso al tempo t ;
- e_{t+1} è il rendimento in eccesso atteso al tempo $t + 1$, ricalcolato in base alle nuove informazioni;
- $\kappa = \frac{p_Z}{\sigma_Q^2}$ è una costante di proporzionalità che collega l'innovazione nei rendimenti attesi all'innovazione nei rendimenti in eccesso;
- $F_{t+1} - \mathbb{E}_t[F_{t+1}]$ è l'innovazione inattesa nel valore fondamentale.

Questa scomposizione ci permette di vedere il rendimento in eccesso Q_{t+1} come la somma di ciò che era atteso (riflesso da e_t) e delle innovazioni inattese.

Autocorrelazione del Rendimento in Eccesso

Dato il procedimento per il rendimento in eccesso come esposto nell'equazione (15), Il coefficiente di autocorrelazione tra Q_t e Q_{t+1} , il quale misura la relazione lineare tra il rendimento in eccesso realizzato in due periodi consecutivi, è dato da:

$$\rho_{Q_t, Q_{t+1}} = \frac{\text{cov}(Q_t, Q_{t+1})}{\sigma_Q^2} = \frac{-(R - \alpha_Z)(1 - \alpha_Z R)p_Z^2 \sigma_Z^2}{(1 - \alpha_Z^2)\sigma_Q^2} \quad (16)$$

Dove:

- $\text{cov}(Q_t, Q_{t+1})$ è la covarianza tra i rendimenti in eccesso nei periodi t e $t + 1$.
- σ_Q^2 è la varianza del rendimento in eccesso Q_t .
- α_Z è il coefficiente di autoregressione che misura la persistenza del processo stocastico Z_t , legato all'avversione al rischio.
- p_Z è un coefficiente che cattura l'impatto delle variazioni dell'avversione al rischio Z_t sul prezzo dell'azione.
- σ_Z^2 è la varianza del processo Z_t , che rappresenta l'incertezza associata all'avversione al rischio.

Chiaramente, $\rho_{Q_t, Q_{t+1}}$ è positivo se $\alpha_Z > 1/R$ e negativo se $\alpha_Z < 1/R$.

3.4.2 Frazioni Ottimali di Azioni Detenute da Investitori di Tipo A e B

Nel modello di Campbell, Grossman e Wang, gli investitori si dividono in due gruppi: tipo A e tipo B. Ogni gruppo ha una diversa avversione al rischio, e di conseguenza detiene una diversa frazione del proprio portafoglio in azioni rischiose. Le frazioni ottimali di azioni detenute da ciascun gruppo sono determinate come segue.

²Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). "Trading Volume and Serial Correlation in Stock Returns." *The Quarterly Journal of Economics*, 108(4), 905-939.

Investitori di Tipo A

La frazione ottimale X_t^a delle azioni detenute dagli investitori di tipo A, con avversione al rischio costante a , è data dalla massimizzazione dell'utilità attesa come è stata definita nell'equazione generale precedentemente sviluppata (4):

$$X_t^a = \frac{\mathbb{E}_t[Q_{t+1}|\mathcal{I}_t]}{a \cdot \text{Var}[Q_{t+1}|\mathcal{I}_t]} = \frac{1}{a} Z_t \quad (17)$$

Dove:

$$\mathbb{E}_t[Q_{t+1}] = \mathbb{E}_t [P_{t+1} + D_{t+1} - RP_t]$$

come è stato descritto nell'eq.(13) e,

Dove:

- $\mathbb{E}_t[Q_{t+1}|\mathcal{I}_t]$ è il rendimento in eccesso atteso condizionato all'informazione disponibile \mathcal{I}_t .
- $\text{Var}[Q_{t+1}|\mathcal{I}_t]$ è la varianza del rendimento in eccesso condizionata all'informazione \mathcal{I}_t .
- a è il coefficiente di avversione al rischio degli investitori di tipo A.
- Z_t è l'avversione al rischio aggregata del mercato al tempo t .

Allo stesso modo, la seconda uguaglianza riprende la definizione di e_t com'è stata definita nei termini di innovazioni e da quest'ultima si può ricavare agevolmente Z_t come:

$$Z_t = \frac{\mathbb{E}_t[Q_{t+1}|\mathcal{I}_t]}{\sigma_Q^2} = \frac{e_t}{\sigma_Q^2}$$

Investitori di Tipo B

Allo stesso modo per gli investitori B, la frazione ottimale X_t^b delle azioni detenute con avversione al rischio b_t , è data da:

$$X_t^b = \frac{\mathbb{E}_t[Q_{t+1}|\mathcal{I}_t]}{b_t \cdot \text{Var}[Q_{t+1}|\mathcal{I}_t]} = \frac{1}{b_t} Z_t \quad (18)$$

Dove:

- b_t è il coefficiente di avversione al rischio degli investitori di tipo B, che può variare nel tempo.

Queste espressioni mostrano che la frazione ottimale di azioni detenute da ciascun tipo di investitore è inversamente proporzionale al loro grado di avversione al rischio. Gli investitori più avversi al rischio (con a o b_t più alti) deterranno una frazione minore di azioni, mentre quelli meno avversi al rischio deterranno una frazione maggiore³.

³Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). "Trading Volume and Serial Correlation in Stock Returns." *The Quarterly Journal of Economics*, 108(4), 905-939.

Differenza nella Frazione di Azioni Detenute: $X_t^a - X_{t-1}^a$

Nel modello, la differenza tra la frazione ottimale di azioni detenute dagli investitori di tipo A al tempo t e al tempo $t - 1$ è espressa come:

$$X_t^a - X_{t-1}^a = \frac{1}{a}(Z_t - Z_{t-1}) \quad (19)$$

Questa equazione rappresenta il cambiamento nella posizione di portafoglio degli investitori di tipo A da un periodo all'altro. La differenza $Z_t - Z_{t-1}$ rappresenta la variazione nell'avversione al rischio aggregata del mercato tra i tempi $t - 1$ e t . Quando Z_t aumenta rispetto a Z_{t-1} , ciò indica che il mercato è diventato più avverso al rischio, spingendo gli investitori di tipo A a ridurre la loro esposizione al rischio (ossia riducono la frazione di azioni detenute). La frazione di azioni che un investitore di tipo A decide di aggiungere o sottrarre dal proprio portafoglio è inversamente proporzionale al suo coefficiente di avversione al rischio a . Investitori con una maggiore avversione al rischio (valori di a più alti) saranno meno reattivi ai cambiamenti in Z_t .

3.4.3 Volume di Trading V_t

Il volume di trading V_t , nel senso più generale, è una misura di quanto gli investitori stanno attivamente comprando o vendendo azioni. Nel modello, il volume di trading è direttamente collegato alla differenza tra X_t^a e X_{t-1}^a ⁴:

$$V_t = \omega |X_t^a - X_{t-1}^a| = \frac{\omega}{a} |Z_t - Z_{t-1}| \quad (20)$$

Dove ω è la proporzione degli investitori di tipo A rispetto al totale degli investitori. Il volume di trading è proporzionale alla variazione nell'avversione al rischio del mercato. Se Z_t cambia significativamente da un periodo all'altro, si osserverà un elevato volume di trading poiché gli investitori di tipo A aggiustano le loro posizioni di portafoglio. Il volume di trading dipende anche dalla frazione di investitori di tipo A nel mercato ω . Se ω è elevato, significa che una parte consistente del mercato è composta da investitori di tipo A, e quindi i cambiamenti in Z_t avranno un impatto maggiore sul volume totale di trading.

Volume Medio di Trading $\mathbb{E}[V_t]$

Il volume medio di trading, ovvero il valore atteso di V_t , è dato da:

$$\mathbb{E}[V_t] = \mathbb{E} \left[\frac{\omega}{a} |Z_t - Z_{t-1}| \right]$$

Calcolo del Valore Atteso Per calcolare il valore atteso del valore assoluto di $Z_t - Z_{t-1}$, consideriamo che Z_t segue un processo autoregressivo di primo ordine (AR(1)):

$$\tilde{Z}_t = \alpha_Z \tilde{Z}_{t-1} + u_{Z,t}$$

Dove $u_{Z,t} \sim \mathcal{N}(0, \sigma_Z^2)$ è una variabile casuale con distribuzione normale. Pertanto, la differenza $Z_t - Z_{t-1}$ segue anch'essa una distribuzione normale con media zero e deviazione

⁴Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). "Trading Volume and Serial Correlation in Stock Returns." *The Quarterly Journal of Economics*, 108(4), 905-939.

standard $\sigma_Z \sqrt{1 - \alpha_Z^2}$. Il valore atteso del valore assoluto di una variabile casuale X con distribuzione normale $X \sim \mathcal{N}(0, \sigma^2)$ è dato da:

$$\mathbb{E}[|X|] = \sigma \sqrt{\frac{2}{\pi}}$$

Applicando questa formula al nostro caso, otteniamo:

$$\mathbb{E}[|Z_t - Z_{t-1}|] = \sigma_Z \sqrt{1 - \alpha_Z^2} \sqrt{\frac{2}{\pi}}$$

Quindi, il volume medio di trading $\mathbb{E}[V_t]$ diventa:

$$\mathbb{E}[V_t] = \frac{2\omega\sigma_Z}{a\sqrt{\pi(1 - \alpha_Z^2)}} \quad (21)$$

Il volume medio di trading è proporzionale alla deviazione standard σ_Z del processo stocastico Z_t . Maggiore è l'incertezza (varianza) associata a Z_t , maggiore sarà il volume medio di trading. La persistenza del processo Z_t , catturata da α_Z , influisce sul volume medio: se α_Z è vicino a 1, Z_t è molto persistente, e la differenza $Z_t - Z_{t-1}$ sarà piccola, riducendo il volume di trading ⁵.

Quest'ultima equazione completa il modello teorico in merito al comportamento dei rendimenti e del volume.

3.4.4 Implicazioni del Modello per la Correlazione tra Volume e Rendimento Seriale

Nel modello proposto da Campbell, gli investitori hanno perfetta informazione riguardo al livello attuale di Z_t , che rappresenta l'avversione al rischio del mercato. Utilizzando il valore di Z_t , gli investitori possono prevedere i futuri rendimenti in eccesso, come mostrato dall'equazione (12). Quando Z_t è alto, gli investitori di tipo B sono altamente avversi al rischio e quindi meno inclini a detenere azioni. Di conseguenza, il prezzo delle azioni deve aggiustarsi verso il basso per incrementare il rendimento in eccesso atteso, incentivando così gli investitori di tipo A a detenere più azioni. Tuttavia, gli econometrici, non osservando direttamente né Z_t né S_t , devono fare inferenze basandosi sui rendimenti in eccesso realizzati e sul volume di trading. Nonostante ciò, queste variabili forniscono comunque informazioni rilevanti sulle dinamiche di Z_t e S_t ⁶.

3.4.5 Predizione dei Ritorni Futuri con Z_t

Nel modello, gli investitori nell'economia hanno informazioni perfette sul livello attuale di Z_t e possono utilizzare Z_t per prevedere i rendimenti futuri. Un basso rendimento dovuto a un calo del prezzo può essere causato sia da un aumento di Z_t che da una bassa realizzazione di S_t , ovvero cattive notizie sui futuri flussi di cassa. I cambiamenti in Z_t generano trading tra gli investitori, mentre le notizie pubbliche sui futuri flussi di cassa ne fanno generalmente meno. Pertanto, bassi rendimenti accompagnati da un alto volume

⁵La derivazione del valore atteso del valore assoluto di una variabile casuale con distribuzione normale è stata ripresa dal lavoro di Borak, Härdle, e Weron (2005): "Stable distributions" in *Computational Statistics*, vol. 2, Springer.

⁶Si potrebbe effettivamente utilizzare un set di informazioni più raffinato che includa dividendi, prezzi e volumi. Questo migliorerebbe le inferenze su Z_t . Tuttavia, per semplicità, utilizziamo solo i rendimenti in eccesso e il volume in questo studio.

di scambio sono probabilmente dovuti a un aumento di Z_t , mentre quelli accompagnati da un basso volume di scambio sono probabilmente dovuti a basse realizzazioni di S_t . Nel caso di un aumento di Z_t , il rendimento in eccesso previsto per il prossimo periodo sarà elevato, mentre nel caso di un basso S_t , non lo sarà. Pertanto, l'autocorrelazione del rendimento azionario dovrebbe diminuire con l'aumento del volume di scambio⁷.

3.4.6 Modello non lineare con le variabili del Volume di Trading e il rendimento

In questa sottosezione vengono utilizzati metodi analitici per sviluppare formalmente l'intuizione alla base del modello teorico sino ad ora approfondito. Nella sottosezione successiva, invece, verranno utilizzati metodi di simulazione per raggiungere un obiettivo simile.

Alla base di tutto si vuole calcolare la componente prevedibile del rendimento in eccesso basata sul rendimento attuale e sul volume: $E[\tilde{Q}_{t+1} | Q_t, V_t] = \sigma_Q^2 E[Z_t | Q_t, V_t]$. Il seguente teorema tiene:

Sotto le assunzioni che vengono fatte sulla struttura dell'economia e sulla distribuzione degli shock, si ha:

$$E[\tilde{Q}_{t+1} | Q_t, V_t] = \phi_Q \tilde{Q}_t - \phi_V \tanh(\theta V_t \tilde{Q}_t) V_t \quad (22)$$

Un'approssimazione quadratica dell'equazione (21) è:

$$E[\tilde{Q}_{t+1} | Q_t] = (\phi_Q - \theta \phi_V V_t^2) \tilde{Q}_t \quad (23)$$

Dove $(\theta \phi_V) > 0$.

Interpretazione del Teorema 2

Il Teorema 2 introduce una relazione tra il rendimento in eccesso futuro atteso \tilde{Q}_{t+1} , il rendimento attuale Q_t e il volume di scambio V_t . Questa relazione cattura sia l'effetto lineare del rendimento attuale che l'effetto non lineare del volume di scambio.

Componente Lineare: La parte lineare $\phi_Q \tilde{Q}_t$ mostra come il rendimento attuale influisce direttamente sul rendimento futuro atteso. Questo suggerisce una persistenza nel rendimento, dove il rendimento attuale è un buon indicatore del rendimento futuro. In particolare:

- Se $V_t = 0$ e $\tilde{Q}_t = 0$, il rendimento in eccesso atteso sarà nullo: $E[\tilde{Q}_{t+1} | Q_t, V_t = 0] = 0$,
- Se $V_t = 0$ ma $\tilde{Q}_t \neq 0$, allora $E[\tilde{Q}_{t+1} | Q_t, V_t = 0] = \phi_Q \tilde{Q}_t$

Componente Non Lineare: La parte non lineare $\phi_V \tanh(\theta V_t \tilde{Q}_t) V_t$ cattura l'effetto del volume di scambio sul rendimento in eccesso. La funzione tangente iperbolica $\tanh(\cdot)$ viene utilizzata per modellare questo effetto non lineare. La tangente iperbolica è una funzione che satura per valori estremi, il che significa che l'impatto del volume di scambio sul rendimento in eccesso non cresce indefinitamente ma tende a stabilizzarsi. Questo è particolarmente utile per modellare situazioni di mercato in cui gli effetti del volume sul prezzo non sono lineari.

⁷Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). "Trading Volume and Serial Correlation in Stock Returns." *The Quarterly Journal of Economics*, 108(4), 905-939.

Approfondimento sulla Tangente Iperbolica

La funzione tangente iperbolica è definita come:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Questa funzione ha un dominio su tutti i numeri reali e un'immagine che varia tra $(-1$ e $1)$. La sua principale proprietà è che per valori molto grandi o molto piccoli di x , la funzione si avvicina asintoticamente a 1 o -1 , rispettivamente. Questo la rende ideale per modellare fenomeni dove l'impatto cresce rapidamente per piccoli valori di x ma poi si stabilizza.

Nell'equazione (22), $\tanh(\theta V_t \tilde{Q}_t)$ permette di rappresentare una relazione non lineare tra il volume di scambio e il rendimento in eccesso, dove l'effetto del volume diventa meno significativo man mano che V_t o \tilde{Q}_t aumentano.

Analisi del Volume e del Rendimento

In questo caso si possono riscontrare due scenari possibili, i quali rispettivamente sono:

- $V_t = 0$: il volume di scambio è zero, quindi non c'è alcun cambiamento nelle preferenze di rischio degli investitori. In questo caso, il rendimento in eccesso atteso è semplicemente $\phi_Q \tilde{Q}_t$.
- $V_t \neq 0$: in questo caso il volume di scambio è diverso da zero e ciò implica che ci sono stati degli shock alle preferenze di rischio. In questo caso, l'effetto non lineare del volume di scambio entra in gioco come descritto dall'equazione (22).

Questa modellazione sofisticata consente di catturare meglio le dinamiche reali del mercato, fornendo una rappresentazione più accurata delle relazioni tra rendimento in eccesso e volume di scambio.

3.4.7 Relazione tra Volume e Rendimento

Nella formulazione del modello, l'equazione (23) può essere riespressa, tenendo conto che il valore atteso del rendimento in eccesso condizionato su Q_t e V_t , anche attraverso la seguente equazione (Per ulteriori dettagli vedi Appendice):

$$\mathbb{E}[\tilde{Q}_{t+1} | Q_t, V_t] = (\phi_0 - \phi_1 \hat{V}) \tilde{Q}_t \quad (24)$$

Dove:

- ϕ_0 rappresenta un parametro che può essere sia positivo che negativo, a seconda delle condizioni del mercato;
- ϕ_1 è positivo e riflette l'impatto del volume ponderato \hat{V} sul rendimento futuro.

Il teorema 2 stabilisce che l'avversione al rischio aggregata Z_t è legata al rendimento azionario ritardato e al rendimento ritardato ponderato per il volume. Possiamo quindi riscrivere l'equazione (24) come:

$$\mathbb{E}[\tilde{Q}_{t+1} | Q_t, V_t] = (\phi_Q - \theta \phi_V V_t^2) Q_t \quad (25)$$

In linea di principio, si può utilizzare l'intera storia dei rendimenti e dei volumi per migliorare le previsioni dei rendimenti futuri. Definendo quindi il set informativo come:

$$\mathcal{F}_t = \{Q_\tau, V_\tau : \tau \leq t\}$$

Il problema econometrico si riduce al calcolo dell'aspettativa condizionata:

$$\mathbb{E}[Q_{t+1} | \mathcal{F}_t] = \sigma_Q^2 \mathbb{E}[Z_t | \mathcal{F}_t]$$

Questa previsione può essere raffinata attraverso approcci iterativi, anche se ciò va oltre lo scopo del presente articolo.

Coerenza con i Risultati Empirici

Il modello fornisce previsioni coerenti con i risultati empirici. In particolare, prevede che un aumento del volume di scambio sia associato a una riduzione dell'autocorrelazione del rendimento, poiché un volume elevato rappresenta un aumento dell'avversione al rischio nel mercato.

3.5 Simulazione Del Modello

Questa sezione analizza i risultati delle simulazioni per verificare la correlazione tra volume e autocorrelazione seriale in un modello con agenti eterogenei. Sebbene l'analisi teorica della sezione precedente indichi che il volume e l'autocorrelazione dovrebbero essere negativamente correlati, non è chiaro se questo effetto sia quantitativamente significativo per valori plausibili dei parametri. Le simulazioni sono quindi condotte per calibrare i parametri del modello e valutare la rilevanza quantitativa di tale relazione.

Descrizione delle Attività Prive di Rischio e Rischiose

Il modello descritto nella Sezione 3.1, con processi normali di guida, è semplice da simulare poiché rappresenta un modello lineare condizionato all'informazione degli investitori. La complessità aumenta quando si considera un insieme informativo più ristretto, che include solo volume e rendimenti. La questione cruciale è la calibrazione dei parametri del modello.

Il tasso di interesse privo di rischio è fissato a $R = 1.01$ su base annua, corrispondente a 0.00004 su base giornaliera, considerando 250 giorni di trading in un anno. Il parametro autoregressivo per i dividendi azionari, α_D , è posto pari a uno, rendendo i dividendi un "random walk". Su base giornaliera, qualsiasi processo realistico dei dividendi avrà un α_D molto vicino a uno. Successivamente, il prezzo delle azioni è normalizzato in modo tale da risultare uguale a uno quando tutti i termini stocastici sono pari a zero e vengono impostati a zero all'inizio delle simulazioni. Questa normalizzazione consente di mantenere il prezzo medio delle azioni vicino a uno durante i periodi di simulazione, anche se il processo di prezzo delle azioni presenta una radice unitaria. La normalizzazione mantiene inoltre la variabilità percentuale del prezzo prossima a quella assoluta, e garantisce che il coefficiente di avversione al rischio assoluto e il coefficiente di avversione relativa al rischio siano simili in presenza di partecipazioni iniziali nelle attività prive di rischio ridotte.

Innovazione di Varianza del Valore Fondamentale F_t

La scelta successiva riguarda la varianza dell'innovazione del valore fondamentale dell'azione F_t , fissata a $\sigma_F^2 = (0.01)^2$, in modo che la deviazione standard del rendimento giornaliero delle azioni, in assenza di variazioni nell'avversione al rischio, sia pari all'1%. Tale valore è leggermente superiore alla media dei dati del dopoguerra. L'equazione (4) fornisce le implicazioni di questa scelta per le varianze dell'innovazione del segnale del dividendo S_t e dell'innovazione del dividendo contemporaneo ϵ_t . Se il segnale di dividendo non è presente, ovvero $\sigma_S^2 = 0$, la varianza implicita dell'innovazione del dividendo ϵ_t è data da:

$$\sigma_\epsilon^2 = \frac{(R-1)^2 \sigma_{F_t}^2}{R^2}$$

Nel caso generale, se $\sigma_S^2 > 0$, la varianza del dividendo contemporaneo è espressa come:

$$\sigma_\epsilon^2 = \frac{(R - \alpha_D)^2 \sigma_F^2}{R^2} - \frac{\sigma_S^2}{R^2}$$

Specificazione dei Due Gruppi di Investitori

In seguito, viene specificata la presenza di due gruppi di investitori. Si assume inizialmente che entrambi i gruppi abbiano lo stesso coefficiente medio di avversione al rischio. Tale coefficiente può essere determinato dall'equazione di domanda delle azioni (18). Quando tutti gli investitori hanno un'avversione al rischio costante a , l'equazione (18) implica che $a = \mathbb{E}[Q_{t+1}]/\text{var}[Q_{t+1}]$. Se il prezzo dell'azione è uguale al valore fondamentale, la varianza di Q_{t+1} è pari a $\sigma_F^2 = (0.01)^2$. Assegnando il valore $a = 3$, si ottiene un valore ragionevole per $\mathbb{E}[Q_{t+1}]$ pari a 0.0003, corrispondente a un tasso annuo del 7.5%. Questa procedura per stimare l'avversione media al rischio è una variante del metodo proposto da Friend e Blume (1975).

Il parametro ω , che rappresenta la proporzione di agenti "market-making", assume un ruolo rilevante. In primo luogo, nell'equazione (20), il volume degli scambi generato da una variazione in Z_t è proporzionale a ω/a . Inoltre, nell'equazione (10), la relazione tra Z_t e l'avversione al rischio dei "liquidity traders" b_t è determinata da ω . Quando ω è ridotto, Z_t varia quasi proporzionalmente a b_t ; mentre per valori più alti di ω , grandi variazioni in b_t producono variazioni minori in Z_t . Nel caso in cui $a = 3$, è necessario adottare un valore molto basso per ω , pari a 0.0005. I risultati delle simulazioni indicano che il turnover è generalmente dello 0.5% o inferiore. In simulazioni esplorative con valori maggiori di ω , le variazioni di Z_t richieste per spiegare l'effetto del volume sull'autocorrelazione producono volumi di scambio eccessivi quando i market-makers presentano un'avversione al rischio pari a $a = 3$. È possibile aumentare la frazione di market-makers incrementando l'avversione al rischio, poiché il volume dipende dal rapporto ω/a . Risultati di simulazione simili vengono ottenuti, ad esempio, con $\omega = 0.005$ e $a = 30$, mantenendo la media di Z_t pari a 3 per allinearsi al rendimento medio delle azioni. La fase più complessa della calibrazione riguarda la specificazione delle dinamiche del processo Z_t . Il processo dovrebbe generare un comportamento realistico del prezzo delle azioni. L'equazione (14) descrive l'innovazione di varianza come $p_Z^2 \sigma_Z^2 + \sigma_F^2$. Tuttavia, il coefficiente p_Z è funzione di σ_Z^2 e degli altri parametri del modello. Nel caso in cui $\sigma_Z^2 = 0$, come mostrato nell'Appendice A, si ottiene $p_Z(0) = -\sigma_F^2/(R - \alpha_Z)$. Il coefficiente $p_Z(0)$ è il valore assunto da p_Z quando Z_t non è influenzato da variazioni dell'avversione al rischio.

3.5.1 Calibrazione del Coefficiente λ e della Varianza σ_Z^2

La calibrazione del modello inizia con la definizione del coefficiente λ , che rappresenta il rapporto tra la deviazione standard delle innovazioni nel prezzo, causate dall'aleatorietà nel processo Z , e la deviazione standard delle innovazioni nel valore fondamentale del titolo, valutato utilizzando $p_Z(0)$. In formula:

$$\lambda \equiv \frac{p_Z(0)\sigma_Z}{\sigma_F} = \frac{\sigma_F\sigma_Z}{R - \alpha_Z}.$$

L'equazione ci permette di risolvere per la varianza σ_Z^2 come segue:

$$\sigma_Z^2 = \frac{\lambda^2(R - \alpha_Z)^2}{\sigma_F^2} \quad (26)$$

Questa equazione può essere sostituita nella condizione che $\sigma_Z^2 \leq \sigma_Z^{*2}$, che garantisce una soluzione reale per il coefficiente p_Z . Possiamo poi riformulare questa condizione in una forma semplice come $\lambda \leq 0.5$. Di conseguenza, solo una limitata variabilità del prezzo può essere generata dai cambiamenti nell'avversione al rischio.

3.5.2 Definizione dei parametri e Analisi dei Risultati

Nelle simulazioni preliminari, il parametro λ è stato variato nell'intervallo consentito da 0 a 0,5, mentre allo stesso tempo veniva variata la persistenza del parametro α_Z nel suo intervallo consentito da 0 a 1. È stata riscontrata una forte relazione negativa tra il volume di trading e l'autocorrelazione del rendimento di primo ordine solo per valori di λ superiori a circa 0,2, e per valori di α_Z inferiori a circa 0,5. Con valori più piccoli di λ , la variazione dell'avversione al rischio non aveva un effetto sufficiente sul comportamento dei prezzi delle azioni per essere rilevata chiaramente, anche con un numero molto elevato di osservazioni. Con valori più elevati di α_Z , i cambiamenti di prezzo causati dalla variazione dell'avversione al rischio diventano ampiamente permanenti, quindi il volume di trading non segnala chiaramente che i movimenti dei prezzi saranno invertiti. Nelle simulazioni finali riportate, sono stati scelti $\lambda = 0.25$ e $\alpha_Z = 0, 0.25$ e 0.5 .

Una volta scelti i valori dei parametri, è stato possibile risolvere per i coefficienti di prezzo p_0 e p_Z i quali giocano un ruolo fondamentale per il calcolo del rendimento in eccesso e quindi del prezzo. L'ultimo passaggio è stato scegliere un dividendo iniziale D_0 che soddisfi il requisito che il prezzo iniziale sia pari a uno. Per questo punto è stato deciso di definire il prezzo P_t pari a $F_t - D_t$ escludendo le variabili p_0 , p_Z e Z_t in quanto al tempo 0 è ragionevole pensare ad un prezzo indipendente da varie aleatorietà in grado di influenzarne il valore. Successivamente, sono state estratte innovazioni normali con le varianze appropriate e creati dati artificiali sui prezzi delle azioni e sul volume di trading. Infine, sono state condotte 1000 simulazioni per ottenere i valori medi dei coefficienti esposti nella tabella 3.1. Le iterazioni sono state condotte con una serie di 5000 periodi (ovvero il numero di osservazioni del campione dal 2004 al 2023 su cui sono state svolte le analisi empiriche).

I risultati delle simulazioni illustrative sono riportati nella Tabella 3.1. La tabella mostra i risultati di regressione per un modello standard di ritorno AR(1) e per un modello che interagisce il rendimento con il volume di trading. Tutti i parametri sono fissati come descritto sopra, tranne il parametro α_Z che descrive la persistenza dei cambiamenti nell'avversione al rischio. Questo parametro è zero nel pannello A della Tabella 3.1, 0.25

nel pannello B e 0.5 nel pannello C. Nel pannello A, viene riscontrato un forte effetto del volume sull'autocorrelazione di primo ordine dei rendimenti. La statistica t sul volume è 2,648, e la significatività dell'autocorrelazione dei rendimenti è descritta nei pannelli successivi.

Tabella 3.1: Simulazioni di Volume e Autocorrelazione di Primo Ordine

$$r_{t+1} = \alpha + (\beta + \gamma_1 V_t) r_t$$

Specificazione	β (t)	γ_1 (t)	R^2
A: $\alpha_Z = 0$			
AR(1)	-0.036 (2.537)		0.001
Volume	-0.033 (2.340)	-11.260 (2.648)	0.003
B: $\alpha_Z = 0.25$			
AR(1)	-0.022 (1.527)		0.001
Volume	-0.020 (1.434)	-10.059 (1.573)	0.002
C: $\alpha_Z = 0.5$			
AR(1)	-0.011 (0.803)		0.000
Volume	-0.011 (0.752)	-9.751 (0.938)	0.001

Commento sui Risultati

Dai risultati ottenuti, è possibile osservare tendenze simili a quelle descritte da Campbell et al. L'aggiunta del volume nelle regressioni produce un incremento nei valori di R^2 , sebbene rimanga relativamente basso in termini assoluti. Per esempio, per $\alpha_Z = 0$, l'aggiunta del volume al modello fa passare il R^2 da 0.001 a 0.003, mentre per $\alpha_Z = 0.25$, si nota un aumento simile. Questo suggerisce che l'interazione tra volume e rendimento contribuisce a spiegare meglio i movimenti del rendimento, ma non in maniera sufficiente per aumentare significativamente la capacità predittiva del modello.

Inoltre, i coefficienti sul volume sono negativi in tutti i casi, in linea con quanto osservato nel testo originale e con quanto confermato dall'analisi empirica, sebbene quest'ultimi erano notevolmente inferiori rispetto a quelli definiti dal modello teorico, indicando una relazione inversa tra il volume e l'autocorrelazione del rendimento. L'entità di questa relazione si riduce all'aumentare della persistenza del parametro α_Z , suggerendo che cambiamenti più persistenti nell'avversione al rischio riducono l'impatto del volume sull'autocorrelazione del rendimento.

Il problema riguardante i movimenti estremi nell'avversione al rischio Z_t è confermato anche nei risultati ottenuti i quali sono esposti nella tabella 3.2. Le deviazioni standard

di Z_t sono elevate, con un valore massimo di 94.88 per $\alpha_Z = 0$, riducendosi a 56.43 per $\alpha_Z = 0.5$, ma comunque rimanendo elevate. Questo suggerisce che il modello implica movimenti estremamente volatili nell'avversione al rischio dei trader di liquidità (b_t), coerentemente con le osservazioni fatte nel testo originale.

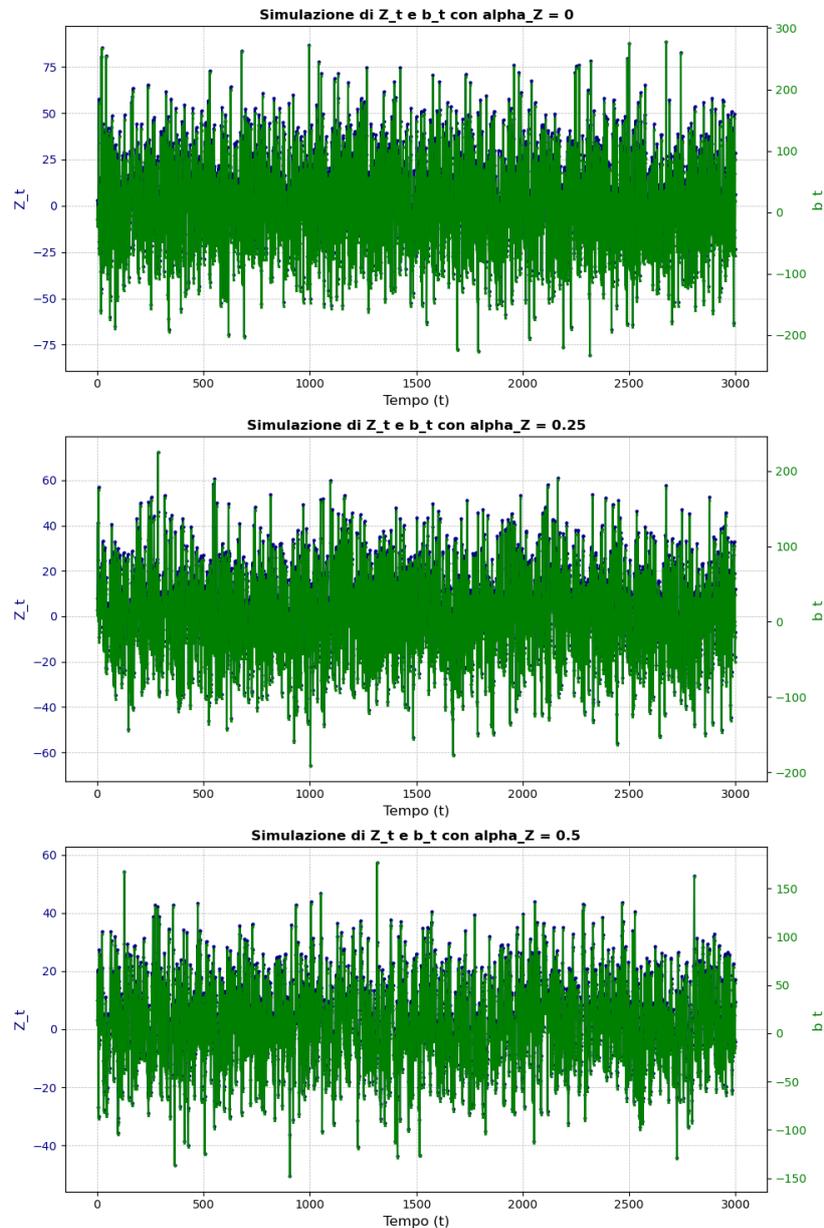


Figura 3.1: Simulazioni di Z_t e b_t con i vari valori di α_Z

In sintesi, mentre l'aggiunta del volume aumenta leggermente la capacità esplicativa del modello, i movimenti estremi nell'avversione al rischio e l'assenza di una forte relazione ad alta frequenza tra volume e rendimento suggeriscono che il modello, così come proposto, ha delle limitazioni nel catturare appieno la dinamica di mercato.

Tabella 3.2: Statistiche di Rendimento, Volume e Z_t per diversi valori di α_Z

Statistiche	R_t	V_t	Z_t
Per $\alpha_Z = 0$			
Media	0.0003	0.005	2.891
Deviazione standard	0.028	0.004	25.287
Min	-0.103	0.000	-89.921
Max	0.106	0.023	94.883
Autocorrelazione lag 1	-0.501	0.225	
Autocorrelazione lag 5	-0.003	0.000	
Per $\alpha_Z = 0.25$			
Media	0.0003	0.003	3.026
Deviazione standard	0.025	0.002	19.610
Min	-0.091	0.000	-68.136
Max	0.095	0.015	74.930
Autocorrelazione lag 1	-0.373	0.123	
Autocorrelazione lag 5	-0.002	0.001	
Per $\alpha_Z = 0.5$			
Media	0.0003	0.002	3.020
Deviazione standard	0.023	0.001	14.720
Min	-0.083	0.000	-52.167
Max	0.088	0.009	56.428
Autocorrelazione lag 1	-0.251	0.055	
Autocorrelazione lag 5	-0.017	0.001	

3.5.3 Implicazioni Empiriche del Modello

Una delle principali difficoltà empiriche del modello risiede nell'effetto che i cambiamenti persistenti in Z_t hanno sui prezzi, senza però generare una forte relazione ad alta frequenza tra il volume di trading e la correlazione seriale. L'interazione tra il volume e il rendimento azionario ritardato (*laggato*) permette di identificare i cambiamenti recenti nel rendimento atteso delle azioni, ma non rappresenta una guida affidabile per il livello attuale del rendimento atteso, soprattutto quando questo segue un processo temporale persistente. D'altra parte, i cambiamenti transitori in Z_t hanno effetti limitati sui prezzi, in quanto piccoli movimenti temporanei possono generare ampie variazioni nei rendimenti attesi. Come evidenziato dall'Equazione (26), all'aumentare della variabilità del parametro di persistenza α_Z , Z_t deve diventare più volatile per qualsiasi impatto sui prezzi, rappresentato dal parametro λ . Di conseguenza, per osservare un forte effetto del volume sulla correlazione seriale, sono necessarie variazioni transitorie molto ampie dell'avversione al rischio. Questo fenomeno riflette la nota difficoltà di spiegare la prevedibilità ad alta frequenza dei rendimenti utilizzando un modello senza frizioni, in cui gli agenti massimizzano l'utilità e sono avversi al rischio. Sebbene il modello consenta di includere un'avversione al rischio eterogenea e variabile nel tempo, esso non risolve completamente questo problema.

Inoltre, il modello mostra una difficoltà empirica legata alla persistenza. È stato rilevato che l'autocorrelazione dei rendimenti azionari dipende da una misura di volume detrendizzato, che risulta essere altamente persistente, con un'autocorrelazione giornaliera di

primo ordine pari a 0.6 e una quinta autocorrelazione che rimane al di sopra di 0.3. Quando si estrae la componente ad alta frequenza del volume, ad esempio mediante un modello a componenti non osservabili o sottraendo una media mobile di alcuni giorni, la relazione tra il volume e l'autocorrelazione risulta significativamente indebolita. Questo risultato è in contrasto con l'implicazione del modello secondo cui il volume segue un processo MA(1) quando l'avversione media al rischio Z_t segue un rumore bianco, o quando si avvicina a un processo MA(1) nel caso in cui l'avversione al rischio segua un processo transitorio AR(1).

Conclusione

Questo studio ha esaminato la relazione tra la volatilità dei mercati azionari e il volume degli scambi, analizzando come variazioni nella domanda degli investitori non informativi e il loro impatto sui market-maker avversi al rischio possano spiegare i movimenti a breve termine nei mercati finanziari. Un risultato chiave dell'analisi mostra che l'autocorrelazione dei rendimenti azionari si riduce significativamente nei giorni di alto volume rispetto a quelli di basso volume, coerentemente con i dati osservati sia nei principali indici azionari che nei singoli titoli.

Tale evidenza suggerisce che l'effetto del volume sia rilevante e non attribuibile a semplici distorsioni causate dalla non sincronia degli scambi. Piuttosto, sembra essere legato a cambiamenti significativi nella liquidità e nelle dinamiche di mercato. La presenza di variazioni casuali nella domanda non informativa e il ruolo giocato dai market-maker nell'assorbire tali variazioni risultano in oscillazioni del volume che hanno un impatto misurabile sui rendimenti. Quando queste variazioni vengono amplificate da movimenti di mercato di breve periodo, i modelli utilizzati indicano che gli agenti avversi al rischio regolano le proprie posizioni in risposta a tali cambiamenti, contribuendo a ridurre l'autocorrelazione nei giorni di elevata attività di scambio.

Nei periodi di crisi finanziarie, come osservato durante la crisi sub-prime o la pandemia da COVID-19, si è rilevata un'evidenza empirica ancora più marcata di tali effetti. In questi periodi, non solo il coefficiente legato al volume di scambi risultava significativamente maggiore, ma anche l' R^2 delle regressioni mostrava un incremento rilevante, segnalando una maggiore capacità del volume di spiegare la variazione dei rendimenti. Questo fenomeno può essere attribuito all'aumento dell'incertezza e del rischio percepito, che induce una maggiore attività di trading da parte degli investitori "liquidi" e una più alta sensibilità dei prezzi ai movimenti del volume.

Infine, attraverso il modello teorico proposto da Campbell, Grossman e Wang sono stati simulati reazioni verosimilmente simili in cui si può considerare che il volume degli scambi avvenga quando variazioni casuali nella domanda di azioni da parte di investitori non informativi vengono assorbite da market-maker avversi al rischio. Quando si permettono grandi variazioni transitorie nella domanda non informativa, il modello riesce a cogliere molte delle caratteristiche osservate nei dati, anche se in questo modello, come in quello precedente esposto da Campbell et al., il modello non cattura con una rilevanza significativa questo comportamento ma ne esalta la struttura logica sottostante.

In sintesi, il modello proposto offre una visione complessa ed efficace della relazione tra volume di scambio, volatilità e autocorrelazione dei rendimenti. Considerando ampie variazioni transitorie nella domanda non informativa e tenendo conto della reattività dei market-maker avversi al rischio, il modello è in grado di spiegare molte delle dinamiche osservate nei dati finanziari, particolarmente durante i periodi di elevata turbolenza economica.

Appendice

In questa appendice sono riportati i risultati di base utilizzati nella trattazione della tesi.

Appendice 3.A: Derivazione del Rendimento in Ecceso Atteso

Per arrivare all'equazione finale completa e dettagliata del valore atteso del rendimento in eccesso $\mathbb{E}[Q_{t+1}|\mathcal{J}_t]$, partiamo dalle informazioni a disposizione, utilizzando la definizione completa di F_t , la scomposizione di Z_t , e considerando che tutte le variabili stocastiche sono distribuite normalmente con media 0.

1. Rendimento in eccesso Q_{t+1}

Il rendimento in eccesso Q_{t+1} è definito come:

$$Q_{t+1} = P_{t+1} + D_{t+1} - RP_t$$

2. Prezzo P_t e P_{t+1}

Il prezzo dell'asset rischioso al tempo t , P_t , è dato da:

$$P_t = F_t - D_t + (p_0 + p_Z Z_t)$$

Il prezzo P_{t+1} sarà definito in modo analogo, ma con le informazioni al tempo $t + 1$:

$$P_{t+1} = F_{t+1} - D_{t+1} + (p_0 + p_Z Z_{t+1})$$

3. Definizione di F_t

Il prezzo fondamentale dell'asset rischioso F_t è determinato come il valore attuale atteso dei flussi di cassa futuri:

$$F_t = \mathbb{E}_t \left[\sum_{s=0}^{\infty} \frac{1}{R^s} D_{t+s} \right] = \frac{\bar{R}\bar{D}}{r} + \frac{R}{R - \alpha_D} \tilde{D}_t + \frac{1}{R - \alpha_D} S_t$$

Dove:

- \bar{R} è il rendimento a lungo termine,
- \bar{D} è il dividendo medio,
- \tilde{D}_t è la componente stocastica del dividendo,
- S_t è il segnale informativo sul dividendo futuro.

4. Espressione per P_t e P_{t+1}

Sostituendo l'espressione di F_t nelle equazioni per P_t e P_{t+1} , otteniamo:

$$P_t = \frac{\bar{R}\bar{D}}{r} + \frac{R}{R - \alpha_D}\tilde{D}_t + \frac{1}{R - \alpha_D}S_t - D_t + p_0 + p_Z Z_t$$

e

$$P_{t+1} = \frac{\bar{R}\bar{D}}{r} + \frac{R}{R - \alpha_D}\tilde{D}_{t+1} + \frac{1}{R - \alpha_D}S_{t+1} - D_{t+1} + p_0 + p_Z Z_{t+1}$$

5. Rendimento in eccesso Q_{t+1}

Ora sostituiamo queste espressioni in $Q_{t+1} = P_{t+1} + D_{t+1} - RP_t$:

$$Q_{t+1} = \left(\frac{\bar{R}\bar{D}}{r} + \frac{R}{R - \alpha_D}\tilde{D}_{t+1} + \frac{1}{R - \alpha_D}S_{t+1} - D_{t+1} + p_0 + p_Z Z_{t+1} \right) + D_{t+1} \\ - R \left(\frac{\bar{R}\bar{D}}{r} + \frac{R}{R - \alpha_D}\tilde{D}_t + \frac{1}{R - \alpha_D}S_t - D_t + p_0 + p_Z Z_t \right)$$

6. Semplificazione dei termini

Eliminiamo i termini comuni come D_{t+1} e $\bar{R}\bar{D}/r$, e raccogliamo i termini in \tilde{D}_t , S_t , e Z_t :

$$Q_{t+1} = \frac{R}{R - \alpha_D}(\tilde{D}_{t+1} - R\tilde{D}_t) + \frac{1}{R - \alpha_D}(S_{t+1} - RS_t) + p_Z(Z_{t+1} - RZ_t) + p_0(1 - R)$$

e

$$\tilde{D}_{t+1} - R\tilde{D}_t = \epsilon_{D,t+1}$$

possiamo riscrivere l'equazione per il rendimento in eccesso come:

$$Q_{t+1} = \frac{R}{R - \alpha_D}(\epsilon_{D,t+1}) + \frac{1}{R - \alpha_D}(S_{t+1} - RS_t) + p_Z(Z_{t+1} - RZ_t) + p_0(1 - R)$$

Sappiamo che:

$$Z_{t+1} = \bar{Z} + \alpha_Z \tilde{Z}_t + u_{Z,t+1}$$

Sostituiamo questa espressione nella formula per Q_{t+1} :

$$Q_{t+1} = p_0(1 - R) + p_Z \left(\bar{Z} + \alpha_Z \tilde{Z}_t + u_{Z,t+1} - R(\bar{Z} + \tilde{Z}_t) \right) + \frac{R}{R - \alpha_D}(\epsilon_{D,t+1}) + \frac{1}{R - \alpha_D}(S_{t+1} - RS_t)$$

Distribuiamo i termini all'interno delle parentesi:

$$Q_{t+1} = p_0(1 - R) + p_Z \left(\bar{Z} + \alpha_Z \tilde{Z}_t + u_{Z,t+1} - R\bar{Z} - R\tilde{Z}_t \right) + \frac{R}{R - \alpha_D}(\epsilon_{D,t+1}) + \frac{1}{R - \alpha_D}(S_{t+1} - RS_t)$$

Raggruppiamo i termini simili:

$$Q_{t+1} = p_0(1 - R) + p_Z \left(\bar{Z}(1 - R) + \tilde{Z}_t(\alpha_Z - R) + u_{Z,t+1} \right) + \frac{R}{R - \alpha_D}(\epsilon_{D,t+1}) + \frac{1}{R - \alpha_D}(S_{t+1} - RS_t)$$

Prendendo il valore atteso e tenendo conto che $u_{Z,t+1}$, $\epsilon_{D,t+1}$ e S_t hanno media zero, allora:

$$E[Q_{t+1}] = p_0(1 - R) + p_Z \left(\bar{Z}(1 - R) + \tilde{Z}_t(\alpha_Z - R) \right)$$

e sapendo che $R = 1 - r$, possiamo riscrivere il valore atteso come:

$$E[Q_{t+1}] = -r(p_0 + p_Z \bar{Z}) + p_Z(\alpha_Z - R)\tilde{Z}_t$$

8.

Verosimilmente, andando a sostituire i parametri che sono stati utilizzati e implementati nel modello simulativo otteniamo:

$$\mathbb{E}[Q_{t+1}|\mathcal{J}_t] = -0.01(-0.0316 + (-0.000141 * 3)) + (-0.000141)(0.25 - 1.01) * 0.23 = 0.0003$$

Questa è l'equazione finale che descrive il rendimento in eccesso atteso.

Appendice 3.B: Derivazione della volatilità attraverso le frazioni degli investitori

1. Soluzione per le ottimali partecipazioni X_t^a e X_t^b

Abbiamo già le espressioni per le partecipazioni ottimali degli investitori di tipo A e B:

$$X_t^a = \frac{E[Q_{t+1}|P_t, D_t, S_t]}{\alpha \text{ var}[Q_{t+1}|P_t, D_t, S_t]} = \frac{1}{\alpha \sigma_Q^2} \left[-r(p_0 + p_Z \bar{Z}) + (\alpha_Z - R)p_Z \tilde{Z}_t \right]$$

$$X_t^b = \frac{E[Q_{t+1}|P_t, D_t, S_t]}{b_t \text{ var}[Q_{t+1}|P_t, D_t, S_t]} = \frac{1}{b_t \sigma_Q^2} \left[-r(p_0 + p_Z \bar{Z}) + (\alpha_Z - R)p_Z \tilde{Z}_t \right]$$

2. Condizione di compensazione del mercato

La condizione di compensazione del mercato stabilisce che la domanda aggregata per l'asset sia pari all'unità:

$$\omega X_t^a + (1 - \omega) X_t^b = 1$$

Sostituendo X_t^a e X_t^b nelle rispettive espressioni, otteniamo:

$$\omega \left(\frac{1}{\alpha \sigma_Q^2} \left[-r(p_0 + p_Z \bar{Z}) + (\alpha_Z - R)p_Z \tilde{Z}_t \right] \right) + (1 - \omega) \left(\frac{1}{b_t \sigma_Q^2} \left[-r(p_0 + p_Z \bar{Z}) + (\alpha_Z - R)p_Z \tilde{Z}_t \right] \right) = 1$$

⁸Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). "Trading Volume and Serial Correlation in Stock Returns." *The Quarterly Journal of Economics*, 108(4), 905-939.

3. Semplificazione dell'equazione di compensazione del mercato

Fattorizzando i termini comuni, abbiamo:

$$\left(\frac{\omega}{\alpha} + \frac{1-\omega}{b_t}\right) \left[-r(p_0 + p_Z \bar{Z}) + (\alpha_Z - R)p_Z \tilde{Z}_t\right] = \sigma_Q^2$$

dove:

$$\left(\frac{\omega}{\alpha} + \frac{1-\omega}{b_t}\right) = \frac{1}{\bar{Z}_t}$$

Questa espressione è importante perché ci fornisce la relazione tra le variabili chiave del modello e σ_Q^2 .

4. Separazione delle componenti di Z_t

Ora separiamo i termini legati alla parte media \bar{Z} e alla parte stocastica \tilde{Z}_t . Ricordiamo che Z_t può essere scomposto in:

$$Z_t = \bar{Z} + \tilde{Z}_t$$

Sostituendo questa scomposizione nell'equazione, otteniamo:

$$\frac{1}{\bar{Z} + \tilde{Z}_t} \left[-r(p_0 + p_Z \bar{Z}) + (\alpha_Z - R)p_Z \tilde{Z}_t\right] = \sigma_Q^2$$

5. Espansione dei termini

Possiamo separare ulteriormente i termini legati alla componente media e alla componente stocastica:

- Per la componente media \bar{Z} :

$$\frac{1}{\bar{Z}} \left[-r(p_0 + p_Z \bar{Z})\right] = \sigma_Q^2$$

- Per la componente stocastica \tilde{Z}_t :

$$\frac{1}{\tilde{Z}_t} \left[(\alpha_Z - R)p_Z \tilde{Z}_t\right] = \sigma_Q^2$$

6. Risoluzione per la componente stocastica

La parte stocastica si semplifica come segue:

$$(\alpha_Z - R)p_Z = \sigma_Q^2$$

7. Risoluzione per la componente media

Moltiplicando entrambi i lati della parte media per \bar{Z} :

$$-r(p_0 + p_Z \bar{Z}) = \sigma_Q^2 \bar{Z}$$

8. Equazioni finali

Infine, otteniamo le due equazioni finali:

$$(\alpha_Z - R)p_Z = \sigma_Q^2$$

e

$$-r(p_0 + p_Z \bar{Z}) = \sigma_Q^2 \bar{Z}$$

Queste due equazioni legano i parametri del modello alla varianza del rendimento in eccesso σ_Q^2 , permettendo di determinare i valori di p_Z e p_0 .

Appendice 3.C: Dimostrazione del Teorema 2

Definiamo $\Delta_t \equiv \left(\frac{\omega}{\alpha}\right) (Z_t - Z_{t-1})$. Pertanto, $V_t = |\Delta_t|$. Definiamo anche $\epsilon_{F,t} = F_t - E_{t-1}[F_t]$, dove $\epsilon_{F,t}$ rappresenta il processo di innovazione associato a F_t .⁹ Allora:

$$Q_{t+1} = p_Z(Z_{t+1} - Z_t) + \epsilon_{F,t+1}$$

Lasciamo che Σ sia la matrice di covarianza di (Q_t, Δ_t) .

Lemma: Dato che Q_{t+1} , Q_t e Δ_t sono congiuntamente normali, possiamo scrivere:

$$E[Q_{t+1}|Q_t, V_t] = \phi_Q Q_t - \phi_V \tanh(\theta V_t \hat{Q}_t) V_t,$$

Dove:

$$\theta = -\frac{\left(\frac{\omega}{\alpha}\right) \sigma_{Q_t, \Delta_t} \hat{Q}_t}{|\Sigma|}$$

$$\phi_Q = \frac{\left(\frac{\omega}{\alpha}\right)^2 \sigma_{Q_t, \Delta_t} \hat{Q}_t}{|\Sigma|} (\sigma_{Q_{t+1}, Q_t} \sigma_{\Delta_t, \Delta_t} - \sigma_{Q_{t+1}, \Delta_t} \sigma_{Q_t, \Delta_t})$$

$$\phi_V = \frac{\left(\frac{\omega}{\alpha}\right)}{|\Sigma|} (\sigma_{Q_{t+1}, Q_t} \sigma_{\Delta_t, \Delta_t} - \sigma_{Q_{t+1}, \Delta_t} \sigma_{Q_t, Q_t})$$

Volume e Correlazione Seriale nei Rendimenti Azionari

Dimostrazione del Lemma: Si veda Wang [1993a]¹⁰. È facile mostrare che:

$$\sigma_{Q_t, Q_t} = \sigma_F^2 + \frac{(1 + R^2 - 2\alpha_Z R)p_Z^2 \sigma_Z^2}{1 - \alpha_Z^2}$$

$$\sigma_{\Delta_t, \Delta_t} = \left(\frac{\omega}{\alpha}\right)^2 \frac{2\sigma_Z^2}{1 + \alpha_Z}$$

$$\sigma_{Q_t, \Delta_t} = \frac{\omega (1 + R)p_Z \sigma_Z^2}{\alpha (1 + \alpha_Z)}$$

⁹Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). "Trading Volume and Serial Correlation in Stock Returns." *The Quarterly Journal of Economics*, 108(4), 905-939.

¹⁰Wang, J. (1993). "A Model of Intertemporal Asset Prices under Asymmetric Information," *Review of Economic Studies*, 60(2), 249-282.

$$\sigma_{Q_{t+1}, Q_t} = \frac{(a_Z - R)(1 - \alpha_Z R)p_Z^2 \sigma_Z^2}{1 - \alpha_Z^2}$$

$$\sigma_{Q_{t+1}, \Delta_t} = \frac{\omega(1 + R)p_Z \sigma_Z^2}{\alpha(1 + \alpha_Z)}$$

Pertanto:

$$\phi_Q = \frac{\left(\frac{\omega}{\alpha}\right)^2 (r + \alpha_Z)(R - \alpha_Z R)p_Z^2 \sigma_Z^2}{(1 + \alpha_Z)(1 - \alpha_Z^2)} > 0$$

$$\phi_V = -\frac{\left(\frac{\omega}{\alpha}\right) (R - \alpha_Z R)p_Z \sigma_Z^2}{1 + \alpha_Z} \left[\sigma_F^2 + \frac{(1 + \alpha_Z R)r p_Z^2 \sigma_Z^2}{1 - \alpha_Z^2} \right]$$

$$\theta = -\frac{\left(\frac{\omega}{\alpha}\right) (1 + R)p_Z \sigma_Z^2}{1 + \alpha_Z}$$

Espressione finale del valore atteso condizionato

Ora possiamo inserire i valori di ϕ_Q , ϕ_V e θ nell'equazione principale:

$$E[Q_{t+1}|Q_t, V_t] = \phi_Q Q_t - \phi_V \tanh(\theta V_t \hat{Q}_t) V_t$$

Pertanto, si conclude con una dimostrazione di come sono stati definiti i vari coefficienti per il rendimento in eccesso condizionato dal rendimento in eccesso al tempo corrente e dal volume di scambio corrente.

Codici Python

In seguito sono illustrati i codici principali e fondamentali che sono stati implementati per la stesura di questo elaborato.

Codici Analisi Empirica

Regressioni con le variabili al tempo T-1 e T-2

```
1 #!/usr/bin/env python
2 # coding: utf-8
3
4 # In[1]:
5
6
7 #Regressioni con T-1
8 import pandas as pd
9 import numpy as np
10 import statsmodels.api as sm
11 from arch import arch_model
12
13 # Caricare i dati dal file Excel
14 file_path = 'S&P500_Hampel.xlsx'
15 df = pd.read_excel(file_path, sheet_name='Foglio1')
16
17 # Converti la colonna delle date in datetime
18 df['Date'] = pd.to_datetime(df['Date'])
19
20 # Calcolo dei rendimenti logaritmici
21 df['Return'] = df['Return'] * 100
22
23 # Calcolo del rendimento quadrato corrente
24 df['Squared_Return'] = df['Return'] ** 2
25
26 # Periodi di campione definiti
27 periods = {
28     'crisi Sub-Prime': ('2005-01-01', '2009-12-31'),
29     'Post crisi subprime': ('2010-01-01', '2014-12-31'),
30     'Ripresa Economica': ('2015-01-01', '2019-12-31'),
31     'Covid-19 e Guerre': ('2020-01-01', '2023-12-31'),
32     'Full Period': ('2004-01-01', '2023-12-31')
33 }
```

```

34
35 # Calcolo della volatilit  condizionata usando GARCH(1,1)
36 garch_model = arch_model(df['Return'].dropna(), vol='Garch', p=1,
    q=1)
37 garch_fit = garch_model.fit(displ='off')
38 df['GARCH_vol'] = garch_fit.conditional_volatility
39
40 results = []
41
42 for period_name, (start_date, end_date) in periods.items():
43     df_period = df[(df['Date'] >= start_date) & (df['Date'] <=
    end_date)].dropna(subset=['Return', 'LogTurnover_filt', '
    GARCH_vol', 'Squared_Return'])
44
45     # Creare le dummies per i giorni della settimana
46     df_period['DayOfWeek'] = df_period['Date'].dt.dayofweek
47     dummies = pd.get_dummies(df_period['DayOfWeek'], prefix='day',
    ).astype(float)
48
49     # Creare le variabili indipendenti moltiplicate per il
    rendimento corrente
50     df_period['Vt_rt'] = df_period['LogTurnover_filt'] *
    df_period['Return']
51     df_period['GARCH_vol_rt'] = df_period['GARCH_vol'] *
    df_period['Return']
52     df_period['Vt_Sq_rt'] = (df_period['LogTurnover_filt'] ** 2)
    * df_period['Return']
53     df_period['Squared_Return_rt'] = df_period['Squared_Return']
    # Rendimento quadrato corrente
54
55     for col in dummies.columns:
56         df_period[col + '_rt'] = dummies[col] * df_period['Return
    ,']
57
58     # Regressione base con volume
59     X_volume = sm.add_constant(df_period[['Vt_rt'] + [col + '_rt'
    for col in dummies.columns]].shift(1).dropna())
60     y_volume = df_period['Return'].loc[X_volume.index]
61     model_volume = sm.OLS(y_volume, X_volume).fit(cov_type='HC3')
62
63     # Regressione con volatilit  condizionata
64     X_volatility = sm.add_constant(df_period[['GARCH_vol_rt'] + [
    col + '_rt' for col in dummies.columns]].shift(1).dropna())
65     y_volatility = df_period['Return'].loc[X_volatility.index]
66     model_volatility = sm.OLS(y_volatility, X_volatility).fit(
    cov_type='HC3')
67
68     # Regressione con volume, volatilit  condizionata e volume
    quadratico + rendimento quadrato
69     X_full = sm.add_constant(df_period[['Vt_rt', 'GARCH_vol_rt',
    'Vt_Sq_rt', 'Squared_Return_rt'] + [col + '_rt' for col in

```

```

dummies.columns]].shift(1).dropna())
70 y_full = df_period['Return'].loc[X_full.index]
71 model_full = sm.OLS(y_full, X_full).fit(cov_type='HC3')
72
73 # Memorizzare i risultati
74 results.append([
75     period_name,
76     model_volume.params.iloc[1], model_volume.bse.iloc[1],
model_volume.rsquared, # Volume
77     model_volatility.params.iloc[1], model_volatility.bse.
iloc[1], model_volatility.rsquared, # Volatility
78     model_full.params.iloc[1], model_full.bse.iloc[1],
model_full.params.iloc[2], model_full.bse.iloc[2], model_full.
params.iloc[3], model_full.bse.iloc[3], model_full.params.iloc
[4], model_full.bse.iloc[4], model_full.rsquared # Volume,
Volatility, Volume^2, Squared_Return
79 ])
80
81 # Creare la tabella dei risultati
82 columns = [
83     'Period',
84     'Volume Coeff', 'Volume SE', 'Volume R ',
85     'Volatility Coeff', 'Volatility SE', 'Volatility R ',
86     'Full Volume Coeff', 'Full Volume SE', 'Full Volatility Coeff
', 'Full Volatility SE', 'Full Volume^2 Coeff', 'Full Volume^2
SE', 'Full Squared_Return Coeff', 'Full Squared_Return SE', '
Full R '
87 ]
88 results_df = pd.DataFrame(results, columns=columns)
89 print(results_df)
90 #-----FINE DELLA REGRESSIONE AL TEMPO T-1
91
92
93 # In[ ]:
94
95
96 #INIZIO DELLA REGRESSIONE AL TEMPO T-2
97 import pandas as pd
98 import numpy as np
99 import statsmodels.api as sm
100 from arch import arch_model
101
102 # Caricare i dati dal file Excel
103 file_path = 'S&P500_Hampel.xlsx'
104 df = pd.read_excel(file_path, sheet_name='Foglio1')
105
106 # Converti la colonna delle date in datetime
107 df['Date'] = pd.to_datetime(df['Date'])
108
109 # Calcolo dei rendimenti logaritmici
110 df['Return'] = df['Return'] * 100

```

```

111
112 # Periodi di campione definiti
113 periods = {
114     'crisi Sub-Prime': ('2005-01-01', '2009-12-31'),
115     'Post crisi subprime': ('2010-01-01', '2014-12-31'),
116     'Ripresa Economica': ('2015-01-01', '2019-12-31'),
117     'Covid-19 e Guerre': ('2020-01-01', '2023-12-31'),
118     'Full Period': ('2004-01-01', '2023-12-31')
119 }
120
121 # Calcolo della volatilit  condizionata usando GARCH(1,1)
122 garch_model = arch_model(df['Return'].dropna(), vol='Garch', p=1,
123                          q=1)
124 garch_fit = garch_model.fit(dispatch='off')
125 df['GARCH_vol'] = garch_fit.conditional_volatility
126
127 results = []
128
129 for period_name, (start_date, end_date) in periods.items():
130     df_period = df[(df['Date'] >= start_date) & (df['Date'] <=
131                 end_date)].dropna(subset=['Return', 'LogTurnover_filt', '
132                 GARCH_vol'])
133
134     # Creare le dummies per i giorni della settimana
135     df_period['DayOfWeek'] = df_period['Date'].dt.dayofweek
136     dummies = pd.get_dummies(df_period['DayOfWeek'], prefix='day',
137                             dtype=float)
138
139     # Creare le variabili indipendenti moltiplicate per il
140     rendimento corrente
141     df_period['Vt_rt'] = df_period['LogTurnover_filt'] *
142     df_period['Return']
143     df_period['GARCH_vol_rt'] = df_period['GARCH_vol'] *
144     df_period['Return']
145     df_period['Vt_Sq_rt'] = (df_period['LogTurnover_filt'] ** 2)
146     * df_period['Return']
147
148     for col in dummies.columns:
149         df_period[col + '_rt'] = dummies[col] * df_period['Return
150     ,]
151
152     # Regressione base con volume
153     X_volume = sm.add_constant(df_period[['Vt_rt'] + [col + '_rt'
154     for col in dummies.columns]].shift(2).dropna())
155     y_volume = df_period['Return'].loc[X_volume.index]
156     model_volume = sm.OLS(y_volume, X_volume).fit(cov_type='HC3')
157
158     # Regressione con volatilit  condizionata
159     X_volatility = sm.add_constant(df_period[['GARCH_vol_rt'] + [
160     col + '_rt' for col in dummies.columns]].shift(2).dropna())
161     y_volatility = df_period['Return'].loc[X_volatility.index]

```

```

151     model_volatility = sm.OLS(y_volatility, X_volatility).fit(
cov_type='HC3')
152
153     # Regressione con volume, volatilita condizionata e volume
quadratico
154     X_full = sm.add_constant(df_period[['Vt_rt', 'GARCH_vol_rt',
'Vt_Sq_rt'] + [col + '_rt' for col in dummies.columns]].shift
(2).dropna())
155     y_full = df_period['Return'].loc[X_full.index]
156     model_full = sm.OLS(y_full, X_full).fit(cov_type='HC3')
157
158     # Memorizzare i risultati
159     results.append([
160         period_name,
161         model_volume.params.iloc[1], model_volume.bse.iloc[1],
model_volume.rsquared, # Volume
162         model_volatility.params.iloc[1], model_volatility.bse.
iloc[1], model_volatility.rsquared, # Volatility
163         model_full.params.iloc[1], model_full.bse.iloc[1],
model_full.params.iloc[2], model_full.bse.iloc[2], model_full.
params.iloc[3], model_full.bse.iloc[3], model_full.rsquared #
164         Volume, Volatility, Volume^2
])
165
166 # Creare la tabella dei risultati
167 columns = [
168     'Period',
169     'Volume Coeff', 'Volume SE', 'Volume R ',
170     'Volatility Coeff', 'Volatility SE', 'Volatility R ',
171     'Full Volume Coeff', 'Full Volume SE', 'Full Volatility Coeff
', 'Full Volatility SE', 'Full Volume^2 Coeff', 'Full Volume^2
SE', 'Full R '
172 ]
173 results_df = pd.DataFrame(results, columns=columns)
174 print(results_df)

```

Listing 3.1: Codice Python per l'analisi dei dati

Regressione con media mobile dei r_t

```

1 #!/usr/bin/env python
2 # coding: utf-8
3
4 # In[ ]:
5
6
7 #REGRESSIONE CON MEDIA MOBILE PASSATA E MEDIA MOBILE INTERAGITA
CON VOLUME T-1
8
9 import pandas as pd
10 import numpy as np
11 import statsmodels.api as sm

```

```

12
13 # Caricare i dati dal file Excel
14 file_path = 'S&P500_Hampel.xlsx'
15 df = pd.read_excel(file_path, sheet_name='Foglio1')
16
17 # Convertire la colonna delle date in datetime
18 df['Date'] = pd.to_datetime(df['Date'])
19
20 # Calcolare i rendimenti logaritmici
21 df['Return'] = df['Return'] * 100
22
23 # Calcolare le medie mobili
24 df['MA_rt'] = df['Return'].rolling(window=5).mean()
25 df['MA_Volume_rt'] = df['MA_rt'] * df['LogTurnover_filt']
26
27 # Periodi di campione definiti
28 periods = {
29     'crisi Sub-Prime': ('2005-01-01', '2009-12-31'),
30     'Post crisi subprime': ('2010-01-01', '2014-12-31'),
31     'Ripresa Economica': ('2015-01-01', '2019-12-31'),
32     'Covid-19 e Guerre': ('2020-01-01', '2023-12-31'),
33     'Full Period': ('2004-01-01', '2023-12-31')
34 }
35
36 results = []
37
38 for period_name, (start_date, end_date) in periods.items():
39     df_period = df[(df['Date'] >= start_date) & (df['Date'] <=
40         end_date)].dropna(subset=['Return', 'MA_rt', 'MA_Volume_rt'])
41
42     # Creare le dummies per i giorni della settimana
43     df_period['DayOfWeek'] = df_period['Date'].dt.dayofweek
44     dummies = pd.get_dummies(df_period['DayOfWeek'], prefix='day',
45         ).astype(float)
46
47     for col in dummies.columns:
48         df_period[col + '_rt'] = dummies[col] * df_period['Return
49         ,']
50
51     # Regressione con la media mobile dei rendimenti passati
52     X_ma = sm.add_constant(df_period[['MA_rt'] + [col + '_rt' for
53         col in dummies.columns]].shift(1).dropna())
54     y_ma = df_period['Return'].loc[X_ma.index]
55     model_ma = sm.OLS(y_ma, X_ma).fit(cov_type='HC3')
56
57     # Regressione con la media mobile dei rendimenti passati
58     # interagita con il volume
59     X_ma_vol = sm.add_constant(df_period[['MA_Volume_rt'] + [col
60         + '_rt' for col in dummies.columns]].shift(1).dropna())
61     y_ma_vol = df_period['Return'].loc[X_ma_vol.index]
62     model_ma_vol = sm.OLS(y_ma_vol, X_ma_vol).fit(cov_type='HC3')

```

```

57
58     # Memorizzare i risultati
59     results.append([
60         period_name,
61         model_ma.params['MA_rt'], model_ma.bse['MA_rt'], model_ma
        .rsquared, # Media mobile dei rendimenti passati
62         model_ma_vol.params['MA_Volume_rt'], model_ma_vol.bse['
        MA_Volume_rt'], model_ma_vol.rsquared # Media mobile dei
        rendimenti passati interagita con il volume
63     ])
64
65 # Creare la tabella dei risultati
66 columns = [
67     'Period',
68     'MA_Return Coeff', 'MA_Return SE', 'MA_Return R ',
69     'MA_Return_Volume Coeff', 'MA_Return_Volume SE', '
        MA_Return_Volume R '
70 ]
71 results_df = pd.DataFrame(results, columns=columns)
72 print(results_df)

```

Listing 3.2: Codice Python per l'analisi dei dati

Regressioni con le componenti non osservate

```

1 #!/usr/bin/env python
2 # coding: utf-8
3
4 # In[1]:
5
6
7 # REGRESSIONI CON COMPONENTE LOCALE NON OSSERVATA
8
9 import pandas as pd
10 import numpy as np
11 import statsmodels.api as sm
12
13 # Caricare i dati dal file Excel
14 file_path = 'S&P500_Hampel.xlsx'
15 df = pd.read_excel(file_path, sheet_name='Foglio1')
16
17 # Converti la colonna delle date in datetime
18 df['Date'] = pd.to_datetime(df['Date'])
19
20 # Calcolo dei rendimenti logaritmici
21 df['Return'] = df['Return'] * 100
22
23 # Modello a componenti non osservate
24 unobserved_model = sm.tsa.UnobservedComponents(df['LogTurnover'],
        dropna(), 'local level')
25 unobserved_result = unobserved_model.fit(dispatch=False)

```

```

26 df.loc[df['LogTurnover'].dropna().index, 'LogTurnover_UC'] =
    unobserved_result.level.smoothed
27 df['LogTurnover_Detrended_UC'] = df['LogTurnover_UC']
28
29 # Periodi di campione definiti
30 periods = {
31     'Crisi Sub-Prime': ('2006-01-01', '2009-12-31'),
32     'Post Crisi Subprime': ('2010-01-01', '2014-12-31'),
33     'Ripresa Economica': ('2015-01-01', '2019-12-31'),
34     'Covid-19 e Guerre': ('2020-01-01', '2023-12-31'),
35     'Full Period': ('2004-01-01', '2023-12-31')
36 }
37
38 results = []
39
40 for period_name, (start_date, end_date) in periods.items():
41     df_period = df[(df['Date'] >= start_date) & (df['Date'] <=
42         end_date)].dropna(subset=['Return', 'LogTurnover_Detrended_UC'
43         ])
44
45     # Creare le dummies per i giorni della settimana, omettendo
46     # una categoria
47     df_period['DayOfWeek'] = df_period['Date'].dt.dayofweek
48     dummies = pd.get_dummies(df_period['DayOfWeek'], prefix='day',
49         , drop_first=True).astype(float)
50
51     # Creare la variabile indipendente moltiplicata per il
52     # rendimento corrente (solo volume)
53     df_period['Vt_rt'] = df_period['LogTurnover_Detrended_UC'] *
54     df_period['Return']
55
56     for col in dummies.columns:
57         df_period[col + '_rt'] = dummies[col] * df_period['Return
58         ',]
59
60     # Regressione solo con il volume
61     X_volume = sm.add_constant(df_period[['Vt_rt'] + [col + '_rt'
62         for col in dummies.columns]]).shift(1).dropna()
63     y_volume = df_period['Return'].loc[X_volume.index]
64     model_volume = sm.OLS(y_volume, X_volume).fit(cov_type='HC3')
65
66     # Memorizzare i risultati
67     results.append([
68         period_name,
69         model_volume.params.iloc[1], model_volume.bse.iloc[1],
70         model_volume.rsquared # Volume
71     ])
72
73 # Creare la tabella dei risultati
74 columns = [
75     'Period',

```

```

67     'Volume Coeff', 'Volume SE', 'Volume R '
68 ]
69 results_df = pd.DataFrame(results, columns=columns)
70 print(results_df)
71
72
73 # In[ ]:
74
75
76 # REGRESSIONE CON MODELLO A COMPONENTI NON OSSERVATE
77
78 import pandas as pd
79 import numpy as np
80 import statsmodels.api as sm
81
82 # Caricare i dati dal file Excel
83 file_path = 'S&P500_Hampel.xlsx'
84 df = pd.read_excel(file_path, sheet_name='Foglio1')
85
86 # Converti la colonna delle date in datetime
87 df['Date'] = pd.to_datetime(df['Date'])
88
89 # Calcolo dei rendimenti logaritmici
90 df['Return'] = df['Return'] * 100
91
92 # Modello a componenti non osservate complesso
93 unobserved_model = sm.tsa.UnobservedComponents(
94     df['LogTurnover'].dropna(),
95     level='local level',
96     trend=True,
97     seasonal=6,
98     cycle=True,
99     irregular=True
100 )
101 unobserved_result = unobserved_model.fit(dispatch=False)
102
103 # Allineare i risultati smoothed con l'indice originale
104 logturnover_uc = pd.Series(unobserved_result.level.smoothed,
105     index=df['LogTurnover'].dropna().index)
106 df['LogTurnover_UC'] = logturnover_uc
107 df['LogTurnover_Detrended_UC'] = df['LogTurnover_UC']
108
109 # Periodi di campione definiti
110 periods = {
111     'Crisi Sub-Prime': ('2006-01-01', '2009-12-31'),
112     'Post Crisi Subprime': ('2010-01-01', '2014-12-31'),
113     'Ripresa Economica': ('2015-01-01', '2019-12-31'),
114     'Covid-19 e Guerre': ('2020-01-01', '2023-12-31'),
115     'Full Period': ('2004-01-01', '2023-12-31')
116 }

```

```

117 results = []
118
119 for period_name, (start_date, end_date) in periods.items():
120     df_period = df[(df['Date'] >= start_date) & (df['Date'] <=
121         end_date)].dropna(subset=['Return', 'LogTurnover_Detrended_UC'
122         ])
123
124     # Creare le dummies per i giorni della settimana, omettendo
125     # una categoria
126     df_period['DayOfWeek'] = df_period['Date'].dt.dayofweek
127     dummies = pd.get_dummies(df_period['DayOfWeek'], prefix='day',
128         , drop_first=True).astype(float)
129
130     # Creare la variabile indipendente moltiplicata per il
131     # rendimento corrente (solo volume)
132     df_period['Vt_rt'] = df_period['LogTurnover_Detrended_UC'] *
133     df_period['Return']
134
135     for col in dummies.columns:
136         df_period[col + '_rt'] = dummies[col] * df_period['Return
137         ,']
138
139     # Regressione base con volume
140     X_volume = sm.add_constant(df_period[['Vt_rt'] + [col + '_rt'
141         for col in dummies.columns]].shift(1).dropna())
142     y_volume = df_period['Return'].loc[X_volume.index]
143     model_volume = sm.OLS(y_volume, X_volume).fit(cov_type='HC3')
144
145     # Memorizzare i risultati
146     results.append([
147         period_name,
148         model_volume.params.iloc[1], model_volume.bse.iloc[1],
149         model_volume.rsquared # Volume
150     ])
151
152 # Creare la tabella dei risultati
153 columns = [
154     'Period',
155     'Volume Coeff', 'Volume SE', 'Volume R '
156 ]
157 results_df = pd.DataFrame(results, columns=columns)
158 print(results_df)

```

Listing 3.3: Codice Python per l'analisi dei dati

Regressioni con il volume detrendizzato con diverse medie mobili

```

1 #!/usr/bin/env python
2 # coding: utf-8
3
4 # In[1]:
5

```

```

6
7 #REGRESSIONE CON VOLUME DETRENDIZZATO CON MEDIA MOBILE
  TRIANGOLARE A 3 MESI
8
9 import pandas as pd
10 import numpy as np
11 import statsmodels.api as sm
12 from arch import arch_model
13 from statsmodels.stats.outliers_influence import
  variance_inflation_factor
14
15 # Caricare i dati dal file Excel
16 file_path = 'S&P500_Hampel.xlsx'
17 df = pd.read_excel(file_path, sheet_name='Foglio1')
18
19 # Convertire la colonna delle date in datetime
20 df['Date'] = pd.to_datetime(df['Date'])
21
22 # Calcolo dei rendimenti logaritmici
23 df['Return'] = df['Return'] * 100
24
25 # Calcola il log volume
26 df['LogTurnover'] = np.log(df['Volume'])
27
28 # Funzione per calcolare la media mobile triangolare
29 def triangular_moving_average(series, window):
30     weights = np.arange(1, window + 1)
31     weights = np.concatenate((weights, weights[::-1][1:])) #
  Creazione del triangolo simmetrico
32     return series.rolling(window=len(weights), center=True).apply
  (lambda x: np.dot(x, weights) / weights.sum(), raw=True)
33
34 # Calcolare la media mobile triangolare con una finestra di 62
  giorni
35 window = 62
36 df['Triangular_Moving_Average'] = triangular_moving_average(df['
  LogTurnover'], window)
37 # Funzione per il filtro di Hampel
38 def hampel_filter(series, window_size, n_sigmas=3):
39     series = series.copy()
40     rolling_median = series.rolling(window=window_size, center=
  True).median()
41     diff = np.abs(series - rolling_median)
42     rolling_std = series.rolling(window=window_size, center=True)
  .std()
43     threshold = n_sigmas * rolling_std
44     outlier_idx = diff > threshold
45     series[outlier_idx] = rolling_median[outlier_idx]
46     return series
47
48 # Detrendizzare utilizzando la media mobile triangolare

```

```

49 df['LogTurnover_Triangular3m_Detrended'] = df['LogTurnover'] - df
    ['Triangular_Moving_Average']
50 df['LogTurnover_Triangular3m_Detrended'] = hampel_filter(df['
    LogTurnover_Triangular3m_Detrended'], window_size=66)
51
52 # Eliminare i valori NaN risultanti dal detrending
53 df_clean = df.dropna()
54
55 # Periodi di campione definiti
56 periods = {
57     'crisi Sub-Prime': ('2005-01-01', '2009-12-31'),
58     'Post crisi subprime': ('2010-01-01', '2014-12-31'),
59     'Ripresa Economica': ('2015-01-01', '2019-12-31'),
60     'Covid-19 e Guerre': ('2020-01-01', '2023-12-31'),
61     'Full Period': ('2004-01-01', '2023-12-31')
62 }
63
64 # Calcolo della volatilit  condizionata usando GARCH(1,1)
65 garch_model = arch_model(df['Return'].dropna(), vol='Garch', p=1,
    q=1)
66 garch_fit = garch_model.fit(dispatch='off')
67 df['GARCH_vol'] = garch_fit.conditional_volatility
68
69 results = []
70
71 for period_name, (start_date, end_date) in periods.items():
72     df_period = df[(df['Date'] >= start_date) & (df['Date'] <=
    end_date)].dropna(subset=['Return', '
    LogTurnover_Triangular3m_Detrended'])
73
74     # Creare le dummies per i giorni della settimana
75     df_period['DayOfWeek'] = df_period['Date'].dt.dayofweek
76     dummies = pd.get_dummies(df_period['DayOfWeek'], prefix='day',
    , drop_first=True).astype(float)
77
78     # Creare le variabili indipendenti moltiplicate per il
    rendimento corrente
79     df_period['Volume_Tri_MA_3M_rt'] = df_period['
    LogTurnover_Triangular3m_Detrended'] * df_period['Return']
80     df_period['GARCH_vol_rt'] = df_period['GARCH_vol'] *
    df_period['Return']
81     df_period['Volume_Tri_MA_3M_Sq_rt'] = (df_period['
    LogTurnover_Triangular3m_Detrended'] ** 2) * df_period['Return
    '
82
83     for col in dummies.columns:
84         df_period[col + '_rt'] = dummies[col] * df_period['Return
    '
85
86     # Regressione base con volume
87     X_volume = sm.add_constant(df_period[['Volume_Tri_MA_3M_rt']

```

```

+ [col + '_rt' for col in dummies.columns]].shift(1).dropna()
88 y_volume = df_period['Return'].loc[X_volume.index]
89 model_volume = sm.OLS(y_volume, X_volume).fit(cov_type='HC3')
90
91 # Regressione con volatilita condizionata
92 X_volatility = sm.add_constant(df_period[['GARCH_vol_rt'] + [
col + '_rt' for col in dummies.columns]].shift(1).dropna())
93 y_volatility = df_period['Return'].loc[X_volatility.index]
94 model_volatility = sm.OLS(y_volatility, X_volatility).fit(
cov_type='HC3')
95
96 # Regressione con volume, volatilita condizionata e volume
quadratico
97 X_full = sm.add_constant(df_period[['Volume_Tri_MA_3M_rt', '
GARCH_vol_rt', 'Volume_Tri_MA_3M_Sq_rt'] + [col + '_rt' for
col in dummies.columns]].shift(1).dropna())
98 y_full = df_period['Return'].loc[X_full.index]
99 model_full = sm.OLS(y_full, X_full).fit(cov_type='HC3')
100
101
102 # Memorizzare i risultati
103 results.append([
104     period_name,
105     model_volume.params.iloc[1], model_volume.bse.iloc[1],
model_volume.rsquared, # Volume
106     model_volatility.params.iloc[1], model_volatility.bse.
iloc[1], model_volatility.rsquared, # Volatility
107     model_full.params.iloc[1], model_full.bse.iloc[1],
model_full.params.iloc[2], model_full.bse.iloc[2], model_full.
params.iloc[3], model_full.bse.iloc[3], model_full.rsquared #
Volume, Volatility, Volume^2
108 ])
109
110 # Creare la tabella dei risultati
111 columns = [
112     'Period',
113     'Volume Coeff', 'Volume SE', 'Volume R ',
114     'Volatility Coeff', 'Volatility SE', 'Volatility R ',
115     'Full Volume Coeff', 'Full Volume SE', 'Full Volatility Coeff
', 'Full Volatility SE', 'Full Volume^2 Coeff', 'Full Volume^2
SE', 'Full R '
116 ]
117 results_df = pd.DataFrame(results, columns=columns)
118 print(results_df)
119
120 #--- FINE CODICE - REGRESSIONE CON VOLUME DETRENDIZZATO CON MEDIA
MOBILE TRIANGOLARE A 3 MESI
121
122
123 # In[2]:
124

```

```

125
126 #REGRESSIONE CON VOLUME DETRENDIZZATO CON MEDIA MOBILE
    TRIANGOLARE A 6 MESI
127
128 import pandas as pd
129 import numpy as np
130 import statsmodels.api as sm
131 from arch import arch_model
132 from statsmodels.stats.outliers_influence import
    variance_inflation_factor
133
134 # Caricare i dati dal file Excel
135 file_path = 'S&P500_Hampel.xlsx'
136 df = pd.read_excel(file_path, sheet_name='Foglio1')
137
138 # Convertire la colonna delle date in datetime
139 df['Date'] = pd.to_datetime(df['Date'])
140
141 # Calcolo dei rendimenti logaritmici
142 df['Return'] = df['Return'] * 100
143
144 # Calcola il log volume
145 df['LogTurnover'] = np.log(df['Volume'])
146
147 # Funzione per calcolare la media mobile triangolare
148 def triangular_moving_average(series, window):
149     weights = np.arange(1, window + 1)
150     weights = np.concatenate((weights, weights[::-1][1:])) #
    Creazione del triangolo simmetrico
151     return series.rolling(window=len(weights), center=True).apply
    (lambda x: np.dot(x, weights) / weights.sum(), raw=True)
152
153 # Calcolare la media mobile triangolare con una finestra di 62
    giorni
154 window = 124
155 df['Triangular_Moving_Average'] = triangular_moving_average(df['
    LogTurnover'], window)
156 # Funzione per il filtro di Hampel
157 def hampel_filter(series, window_size, n_sigmas=3):
158     series = series.copy()
159     rolling_median = series.rolling(window=window_size, center=
    True).median()
160     diff = np.abs(series - rolling_median)
161     rolling_std = series.rolling(window=window_size, center=True)
    .std()
162     threshold = n_sigmas * rolling_std
163     outlier_idx = diff > threshold
164     series[outlier_idx] = rolling_median[outlier_idx]
165     return series
166
167 # Detrendizzare utilizzando la media mobile triangolare

```

```

168 df['LogTurnover_Triangular3m_Detrended'] = df['LogTurnover'] - df
    ['Triangular_Moving_Average']
169 df['LogTurnover_Triangular3m_Detrended'] = hampel_filter(df['
    LogTurnover_Triangular3m_Detrended'], window_size=66)
170
171 # Eliminare i valori NaN risultanti dal detrending
172 df_clean = df.dropna()
173
174 # Periodi di campione definiti
175 periods = {
176     'crisi Sub-Prime': ('2005-01-01', '2009-12-31'),
177     'Post crisi subprime': ('2010-01-01', '2014-12-31'),
178     'Ripresa Economica': ('2015-01-01', '2019-12-31'),
179     'Covid-19 e Guerre': ('2020-01-01', '2023-12-31'),
180     'Full Period': ('2004-01-01', '2023-12-31')
181 }
182
183 # Calcolo della volatilit  condizionata usando GARCH(1,1)
184 garch_model = arch_model(df['Return'].dropna(), vol='Garch', p=1,
    q=1)
185 garch_fit = garch_model.fit(dispatch='off')
186 df['GARCH_vol'] = garch_fit.conditional_volatility
187
188 results = []
189
190 for period_name, (start_date, end_date) in periods.items():
191     df_period = df[(df['Date'] >= start_date) & (df['Date'] <=
    end_date)].dropna(subset=['Return', '
    LogTurnover_Triangular3m_Detrended'])
192
193     # Creare le dummies per i giorni della settimana
194     df_period['DayOfWeek'] = df_period['Date'].dt.dayofweek
195     dummies = pd.get_dummies(df_period['DayOfWeek'], prefix='day',
    , drop_first=True).astype(float)
196
197     # Creare le variabili indipendenti moltiplicate per il
    rendimento corrente
198     df_period['Volume_Tri_MA_3M_rt'] = df_period['
    LogTurnover_Triangular3m_Detrended'] * df_period['Return']
199     df_period['GARCH_vol_rt'] = df_period['GARCH_vol'] *
    df_period['Return']
200     df_period['Volume_Tri_MA_3M_Sq_rt'] = (df_period['
    LogTurnover_Triangular3m_Detrended'] ** 2) * df_period['Return
    '
201
202     for col in dummies.columns:
203         df_period[col + '_rt'] = dummies[col] * df_period['Return
    '
204
205     # Regressione base con volume
206     X_volume = sm.add_constant(df_period[['Volume_Tri_MA_3M_rt']

```

```

+ [col + '_rt' for col in dummies.columns]].shift(1).dropna())
207 y_volume = df_period['Return'].loc[X_volume.index]
208 model_volume = sm.OLS(y_volume, X_volume).fit(cov_type='HC3')
209
210 # Regressione con volatilita condizionata
211 X_volatility = sm.add_constant(df_period[['GARCH_vol_rt'] + [
col + '_rt' for col in dummies.columns]].shift(1).dropna())
212 y_volatility = df_period['Return'].loc[X_volatility.index]
213 model_volatility = sm.OLS(y_volatility, X_volatility).fit(
cov_type='HC3')
214
215 # Regressione con volume, volatilita condizionata e volume
quadratico
216 X_full = sm.add_constant(df_period[['Volume_Tri_MA_3M_rt', '
GARCH_vol_rt', 'Volume_Tri_MA_3M_Sq_rt'] + [col + '_rt' for
col in dummies.columns]].shift(1).dropna())
217 y_full = df_period['Return'].loc[X_full.index]
218 model_full = sm.OLS(y_full, X_full).fit(cov_type='HC3')
219
220
221 # Memorizzare i risultati
222 results.append([
223     period_name,
224     model_volume.params.iloc[1], model_volume.bse.iloc[1],
model_volume.rsquared, # Volume
225     model_volatility.params.iloc[1], model_volatility.bse.
iloc[1], model_volatility.rsquared, # Volatility
226     model_full.params.iloc[1], model_full.bse.iloc[1],
model_full.params.iloc[2], model_full.bse.iloc[2], model_full.
params.iloc[3], model_full.bse.iloc[3], model_full.rsquared #
Volume, Volatility, Volume^2
227 ])
228
229 # Creare la tabella dei risultati
230 columns = [
231     'Period',
232     'Volume Coeff', 'Volume SE', 'Volume R ',
233     'Volatility Coeff', 'Volatility SE', 'Volatility R ',
234     'Full Volume Coeff', 'Full Volume SE', 'Full Volatility Coeff
', 'Full Volatility SE', 'Full Volume^2 Coeff', 'Full Volume^2
SE', 'Full R '
235 ]
236 results_df = pd.DataFrame(results, columns=columns)
237 print(results_df)
238
239 #----- FINE CODICE - REGRESSIONE CON VOLUME DETRENDIZZATO CON
MEDIA MOBILE TRIANGOLARE A 6 MESI
240
241
242 # In[ ]:
243

```

```

244
245 #REGRESSIONE CON VOLUME DETRENDIZZATO CON MEDIA MOBILE
    TRIANGOLARE A 3 ANNI
246
247 import pandas as pd
248 import numpy as np
249 import statsmodels.api as sm
250 from arch import arch_model
251 from statsmodels.stats.outliers_influence import
    variance_inflation_factor
252
253 # Caricare i dati dal file Excel
254 file_path = 'S&P500_Hampel.xlsx'
255 df = pd.read_excel(file_path, sheet_name='Foglio1')
256
257 # Convertire la colonna delle date in datetime
258 df['Date'] = pd.to_datetime(df['Date'])
259
260 # Calcolo dei rendimenti logaritmici
261 df['Return'] = df['Return'] * 100
262
263 # Calcola il log volume
264 df['LogTurnover'] = np.log(df['Volume'])
265
266 # Funzione per calcolare la media mobile triangolare
267 def triangular_moving_average(series, window):
268     weights = np.arange(1, window + 1)
269     weights = np.concatenate((weights, weights[::-1][1:])) #
    Creazione del triangolo simmetrico
270     return series.rolling(window=len(weights), center=True).apply
    (lambda x: np.dot(x, weights) / weights.sum(), raw=True)
271
272 # Calcolare la media mobile triangolare con una finestra di 756
    giorni
273 window = 756
274 df['Triangular_Moving_Average'] = triangular_moving_average(df['
    LogTurnover'], window)
275 # Funzione per il filtro di Hampel
276 def hampel_filter(series, window_size, n_sigmas=3):
277     series = series.copy()
278     rolling_median = series.rolling(window=window_size, center=
    True).median()
279     diff = np.abs(series - rolling_median)
280     rolling_std = series.rolling(window=window_size, center=True)
    .std()
281     threshold = n_sigmas * rolling_std
282     outlier_idx = diff > threshold
283     series[outlier_idx] = rolling_median[outlier_idx]
284     return series
285
286 # Detrendizzare utilizzando la media mobile triangolare

```

```

287 df['LogTurnover_Triangular3m_Detrended'] = df['LogTurnover'] - df
    ['Triangular_Moving_Average']
288 df['LogTurnover_Triangular3m_Detrended'] = hampel_filter(df['
    LogTurnover_Triangular3m_Detrended'], window_size=66)
289
290 # Eliminare i valori NaN risultanti dal detrending
291 df_clean = df.dropna()
292
293 # Periodi di campione definiti
294 periods = {
295     'crisi Sub-Prime': ('2005-01-01', '2009-12-31'),
296     'Post crisi subprime': ('2010-01-01', '2014-12-31'),
297     'Ripresa Economica': ('2015-01-01', '2019-12-31'),
298     'Covid-19 e Guerre': ('2020-01-01', '2023-12-31'),
299     'Full Period': ('2004-01-01', '2023-12-31')
300 }
301
302 # Calcolo della volatilit  condizionata usando GARCH(1,1)
303 garch_model = arch_model(df['Return'].dropna(), vol='Garch', p=1,
    q=1)
304 garch_fit = garch_model.fit(dispatch='off')
305 df['GARCH_vol'] = garch_fit.conditional_volatility
306
307 results = []
308
309 for period_name, (start_date, end_date) in periods.items():
310     df_period = df[(df['Date'] >= start_date) & (df['Date'] <=
    end_date)].dropna(subset=['Return', '
    LogTurnover_Triangular3m_Detrended'])
311
312     # Creare le dummies per i giorni della settimana
313     df_period['DayOfWeek'] = df_period['Date'].dt.dayofweek
314     dummies = pd.get_dummies(df_period['DayOfWeek'], prefix='day',
    , drop_first=True).astype(float)
315
316     # Creare le variabili indipendenti moltiplicate per il
    rendimento corrente
317     df_period['Volume_Tri_MA_3M_rt'] = df_period['
    LogTurnover_Triangular3m_Detrended'] * df_period['Return']
318     df_period['GARCH_vol_rt'] = df_period['GARCH_vol'] *
    df_period['Return']
319     df_period['Volume_Tri_MA_3M_Sq_rt'] = (df_period['
    LogTurnover_Triangular3m_Detrended'] ** 2) * df_period['Return
    '
320
321     for col in dummies.columns:
322         df_period[col + '_rt'] = dummies[col] * df_period['Return
    '
323
324     # Regressione base con volume
325     X_volume = sm.add_constant(df_period[['Volume_Tri_MA_3M_rt']

```

```

+ [col + '_rt' for col in dummies.columns]].shift(1).dropna()
326 y_volume = df_period['Return'].loc[X_volume.index]
327 model_volume = sm.OLS(y_volume, X_volume).fit(cov_type='HC3')
328
329 # Regressione con volatilita condizionata
330 X_volatility = sm.add_constant(df_period[['GARCH_vol_rt'] + [
col + '_rt' for col in dummies.columns]].shift(1).dropna())
331 y_volatility = df_period['Return'].loc[X_volatility.index]
332 model_volatility = sm.OLS(y_volatility, X_volatility).fit(
cov_type='HC3')
333
334 # Regressione con volume, volatilita condizionata e volume
quadratico
335 X_full = sm.add_constant(df_period[['Volume_Tri_MA_3M_rt', '
GARCH_vol_rt', 'Volume_Tri_MA_3M_Sq_rt'] + [col + '_rt' for
col in dummies.columns]].shift(1).dropna())
336 y_full = df_period['Return'].loc[X_full.index]
337 model_full = sm.OLS(y_full, X_full).fit(cov_type='HC3')
338
339
340 # Memorizzare i risultati
341 results.append([
342     period_name,
343     model_volume.params.iloc[1], model_volume.bse.iloc[1],
model_volume.rsquared, # Volume
344     model_volatility.params.iloc[1], model_volatility.bse.
iloc[1], model_volatility.rsquared, # Volatility
345     model_full.params.iloc[1], model_full.bse.iloc[1],
model_full.params.iloc[2], model_full.bse.iloc[2], model_full.
params.iloc[3], model_full.bse.iloc[3], model_full.rsquared #
Volume, Volatility, Volume^2
346 ])
347
348 # Creare la tabella dei risultati
349 columns = [
350     'Period',
351     'Volume Coeff', 'Volume SE', 'Volume R ',
352     'Volatility Coeff', 'Volatility SE', 'Volatility R ',
353     'Full Volume Coeff', 'Full Volume SE', 'Full Volatility Coeff
', 'Full Volatility SE', 'Full Volume^2 Coeff', 'Full Volume^2
SE', 'Full R '
354 ]
355 results_df = pd.DataFrame(results, columns=columns)
356 print(results_df)

```

Listing 3.4: Codice Python per l'analisi dei dati

Regressione con l'integrazione del rendimento quadratico

```

1 #!/usr/bin/env python
2 # coding: utf-8
3

```

```

4 # In[ ]:
5
6
7 # REGRESSIONE CON RENDIMENTO QUADRATICO
8 import pandas as pd
9 import numpy as np
10 import statsmodels.api as sm
11 from arch import arch_model
12
13 # Caricare i dati dal file Excel
14 file_path = 'S&P500_Hampel.xlsx'
15 df = pd.read_excel(file_path, sheet_name='Foglio1')
16
17 # Converti la colonna delle date in datetime
18 df['Date'] = pd.to_datetime(df['Date'])
19
20 # Calcolo dei rendimenti logaritmici
21 df['Return'] = df['Return'] * 100
22
23 # Calcolo del rendimento quadrato corrente
24 df['Squared_Return'] = df['Return'] ** 2
25
26 # Periodi di campione definiti
27 periods = {
28     'crisi Sub-Prime': ('2006-01-01', '2009-12-31'),
29     'Post crisi subprime': ('2010-01-01', '2014-12-31'),
30     'Ripresa Economica': ('2015-01-01', '2019-12-31'),
31     'Covid-19 e Guerre': ('2020-01-01', '2023-12-31'),
32     'Full Period': ('2004-01-01', '2023-12-31')
33 }
34
35 # Calcolo della volatilit  condizionata usando GARCH(1,1)
36 garch_model = arch_model(df['Return'].dropna(), vol='Garch', p=1,
37     q=1)
38 garch_fit = garch_model.fit(dispatch='off')
39 df['GARCH_vol'] = garch_fit.conditional_volatility
40
41 results = []
42
43 for period_name, (start_date, end_date) in periods.items():
44     df_period = df[(df['Date'] >= start_date) & (df['Date'] <=
45         end_date)].dropna(subset=['Return', 'LogTurnover_filt', '
46         GARCH_vol', 'Squared_Return'])
47
48     # Creare le dummies per i giorni della settimana
49     df_period['DayOfWeek'] = df_period['Date'].dt.dayofweek
50     dummies = pd.get_dummies(df_period['DayOfWeek'], prefix='day',
51         ).astype(float)
52
53     # Creare le variabili indipendenti moltiplicate per il
54     rendimento corrente

```

```

50 df_period['Vt_rt'] = df_period['LogTurnover_filt'] *
df_period['Return']
51 df_period['GARCH_vol_rt'] = df_period['GARCH_vol'] *
df_period['Return']
52 df_period['Vt_Sq_rt'] = (df_period['LogTurnover_filt'] ** 2)
* df_period['Return']
53 df_period['Squared_Return_rt'] = df_period['Squared_Return']
# Rendimento quadrato corrente
54
55 for col in dummies.columns:
56     df_period[col + '_rt'] = dummies[col] * df_period['Return
,']
57
58 # Regressione base con volume
59 X_volume = sm.add_constant(df_period[['Vt_rt'] + [col + '_rt'
for col in dummies.columns]].shift(1).dropna())
60 y_volume = df_period['Return'].loc[X_volume.index]
61 model_volume = sm.OLS(y_volume, X_volume).fit(cov_type='HC3')
62
63 # Regressione con volatilita condizionata
64 X_volatility = sm.add_constant(df_period[['GARCH_vol_rt'] + [
col + '_rt' for col in dummies.columns]].shift(1).dropna())
65 y_volatility = df_period['Return'].loc[X_volatility.index]
66 model_volatility = sm.OLS(y_volatility, X_volatility).fit(
cov_type='HC3')
67
68 # Regressione con volume, volatilita condizionata e volume
quadratico + rendimento quadrato
69 X_full = sm.add_constant(df_period[['Vt_rt', 'GARCH_vol_rt',
'Vt_Sq_rt', 'Squared_Return_rt'] + [col + '_rt' for col in
dummies.columns]].shift(1).dropna())
70 y_full = df_period['Return'].loc[X_full.index]
71 model_full = sm.OLS(y_full, X_full).fit(cov_type='HC3')
72
73 # Memorizzare i risultati
74 results.append([
75     period_name,
76     model_volume.params.iloc[1], model_volume.bse.iloc[1],
model_volume.rsquared, # Volume
77     model_volatility.params.iloc[1], model_volatility.bse.
iloc[1], model_volatility.rsquared, # Volatility
78     model_full.params.iloc[1], model_full.bse.iloc[1],
model_full.params.iloc[2], model_full.bse.iloc[2], model_full.
params.iloc[3], model_full.bse.iloc[3], model_full.params.iloc
[4], model_full.bse.iloc[4], model_full.rsquared # Volume,
Volatility, Volume^2, Squared_Return
79     ])
80
81 # Creare la tabella dei risultati
82 columns = [
83     'Period',

```

```

84     'Volume Coeff', 'Volume SE', 'Volume R ',
85     'Volatility Coeff', 'Volatility SE', 'Volatility R ',
86     'Full Volume Coeff', 'Full Volume SE', 'Full Volatility Coeff
', 'Full Volatility SE', 'Full Volume^2 Coeff', 'Full Volume^2
SE', 'Full Squared_Return Coeff', 'Full Squared_Return SE', '
Full R '
87 ]
88 results_df = pd.DataFrame(results, columns=columns)
89 print(results_df)

```

Listing 3.5: Codice Python per l'analisi dei dati

Regressioni multiple con le maggiori 30 società

```

1  #!/usr/bin/env python
2  # coding: utf-8
3
4  # In[1]:
5
6
7  # REGRESSIONI MULTIPLE CON 30 SOCIETA'
8  import pandas as pd
9  import numpy as np
10
11 # Caricare i dati dal file Excel
12 file_path = '30top_Hampel.xlsx'
13 df = pd.read_excel(file_path, sheet_name='Sheet1')
14
15 # Funzione per calcolare la media mobile triangolare
16 def triangular_moving_average(series, window):
17     weights = np.arange(1, window + 1)
18     weights = np.concatenate((weights, weights[::-1][1:])) #
19     Creazione del triangolo simmetrico
20     return series.rolling(window=len(weights), center=True).apply
21     (lambda x: np.dot(x, weights) / weights.sum(), raw=True)
22
23 # Funzione per il filtro di Hampel
24 def hampel_filter(series, window_size, n_sigmas=3):
25     series = series.copy()
26     rolling_median = series.rolling(window=window_size, center=
27     True).median()
28     diff = np.abs(series - rolling_median)
29     rolling_std = series.rolling(window=window_size, center=True)
30     .std()
31     threshold = n_sigmas * rolling_std
32     outlier_idx = diff > threshold
33     series[outlier_idx] = rolling_median[outlier_idx]
34     return series
35
36 # Calcolare il log volume per ciascuna azienda, detrendizzare e
37     applicare il filtro di Hampel
38     window = 250

```

```

34 hampel_window_size = 66 # Finestra per il filtro di Hampel
35
36 for col in df.columns:
37     if 'Volume' in col:
38         company_name = col.split('_')[0]
39         df[f'{company_name}_LogTurnover'] = np.log(df[col])
40         df[f'{company_name}_Triangular_Moving_Average'] =
triangular_moving_average(df[f'{company_name}_LogTurnover'],
window)
41         df[f'{company_name}_LogTurnover_Triangular_Detrended'] =
df[f'{company_name}_LogTurnover'] - df[f'{company_name}
_Triangular_Moving_Average']
42         df[f'{company_name}
_LogTurnover_Triangular_Detrended_Hampel'] = hampel_filter(df[
f'{company_name}_LogTurnover_Triangular_Detrended'],
hampel_window_size)
43
44 # Eliminare i valori NaN risultanti dal detrending
45 df_clean = df.dropna()
46
47 # Calcolare l'indice equal-weighted
48 price_cols = [col for col in df_clean.columns if 'Price' in col]
49 df_clean['Equal_Weighted_Index'] = df_clean[price_cols].mean(axis
=1)
50 df_clean['EW_Return'] = np.log(df_clean['Equal_Weighted_Index'] /
df_clean['Equal_Weighted_Index'].shift(1))
51
52 # Calcolare il log turnover per l'indice equal-weighted
53 volume_cols = [col for col in df_clean.columns if 'Volume' in col
]
54 df_clean['Equal_Weighted_Volume'] = df_clean[volume_cols].sum(
axis=1)
55 df_clean['EW_LogTurnover'] = np.log(df_clean['
Equal_Weighted_Volume'])
56
57 # Calcolare la media mobile triangolare per il log turnover dell'
indice equal-weighted
58 df_clean['EW_Triangular_Moving_Average'] =
triangular_moving_average(df_clean['EW_LogTurnover'], window)
59 df_clean['EW_LogTurnover_Triangular_Detrended'] = df_clean['
EW_LogTurnover'] - df_clean['EW_Triangular_Moving_Average']
60
61 # Applicare il filtro di Hampel al log turnover detrendizzato
dell'indice equal-weighted
62 df_clean['EW_LogTurnover_Triangular_Detrended_Hampel'] =
hampel_filter(df_clean['EW_LogTurnover_Triangular_Detrended'],
hampel_window_size)
63
64 # Salvare il dataframe pulito in un nuovo file Excel
65 output_path = '30top_with_LogTurnover_Hampel.xlsx'
66 df_clean.to_excel(output_path, index=False)

```

```

67
68 # Visualizzare le prime righe del dataframe pulito per confermare
    i cambiamenti
69 print(df_clean.head())
70
71
72 # In[ ]:
73
74
75 import pandas as pd
76 import numpy as np
77 import statsmodels.api as sm
78
79 # Caricare i dati
80 file_path = '30top_with_LogTurnover_Hampel.xlsx'
81 df = pd.read_excel(file_path, sheet_name='Sheet1')
82
83 # Calcolare i rendimenti per ciascuna azienda
84 for col in df.columns:
85     if 'Price' in col:
86         company = col.split('_')[0]
87         df[f'{company}_Return'] = np.log(df[col] / df[col].shift
    (1))
88
89 # Calcolare l'indice equal-weighted
90 price_cols = [col for col in df.columns if 'Price' in col]
91 df['Equal_Weighted_Index'] = df[price_cols].mean(axis=1)
92 df['EW_Return'] = np.log(df['Equal_Weighted_Index'] / df['
    Equal_Weighted_Index'].shift(1))
93
94 # Utilizzare il log turnover gi detrendizzato e filtrato
95 volume_cols = [col for col in df.columns if 'Volume' in col]
96 df['Equal_Weighted_Index_Volume'] = df[volume_cols].mean(axis=1)
97 df['Equal_Weighted_Index_LogTurnover_Triangular_Detrended_Hampel'
    ] = np.log(df['Equal_Weighted_Index_Volume'])
98
99 # Creare dummies per i giorni della settimana
100 df['DayOfWeek'] = df['Date'].dt.dayofweek
101 dummies = pd.get_dummies(df['DayOfWeek'], prefix='day').astype(
    float)
102
103 results = []
104
105 # Periodi di campione definiti
106 periods = {
107     'crisi Sub-Prime': ('2006-01-01', '2009-12-31'),
108     'Post crisi subprime': ('2010-01-01', '2012-12-31'),
109     'Ripresa Economica': ('2013-01-01', '2019-12-31'),
110     'Covid-19 e Guerre': ('2020-01-01', '2023-12-31'),
111     'Full Period': ('2004-01-01', '2023-12-31')
112 }

```

```

113
114 for period_name, (start_date, end_date) in periods.items():
115     df_period = df[(df['Date'] >= start_date) & (df['Date'] <=
116         end_date)].dropna(subset=['EW_Return'])
117
118     # Creare la variabile Vt_rt per l'indice equal-weighted
119     df_period['EW_Vt_rt'] = df_period['
120     Equal_Weighted_Index_LogTurnover_Triangular_Detrended_Hampel']
121     * df_period['EW_Return']
122
123     # Aggiungere le dummies moltiplicate per il rendimento
124     for col in dummies.columns:
125         df_period[f'{col}_rt'] = dummies[col] * df_period['
126         EW_Return']
127
128     # Verificare che i dati non siano vuoti prima di eseguire la
129     regressione
130     X_ew = sm.add_constant(df_period[['EW_Vt_rt'] + [f'{col}_rt'
131     for col in dummies.columns]].shift(1)).dropna()
132     y_ew = df_period['EW_Return'].dropna().loc[X_ew.index]
133
134     if not X_ew.empty and not y_ew.empty:
135         model_ew = sm.OLS(y_ew, X_ew).fit(cov_type='HC3')
136         gamma_ew = model_ew.params['EW_Vt_rt']
137         se_ew = model_ew.bse['EW_Vt_rt']
138         r2_ew = model_ew.rsquared
139     else:
140         gamma_ew, se_ew, r2_ew = np.nan, np.nan, np.nan
141
142     # Regressione pooled
143     pooled_results = []
144     pooled_r2 = []
145     for col in price_cols:
146         company = col.split('_')[0]
147         df_period[f'{company}_Vt_rt'] = df_period[f'{company}
148         _LogTurnover_Triangular_Detrended_Hampel'] * df_period[f'{
149         company}_Return']
150
151         X_pooled = sm.add_constant(df_period[[f'{company}_Vt_rt']
152         + [f'{col}_rt' for col in dummies.columns]].shift(1)).dropna
153         ()
154         y_pooled = df_period[f'{company}_Return'].dropna().loc[
155         X_pooled.index]
156
157         if not X_pooled.empty and not y_pooled.empty:
158             model_pooled = sm.OLS(y_pooled, X_pooled).fit(
159             cov_type='HC3')
160             pooled_results.append(model_pooled.params[1])
161             pooled_r2.append(model_pooled.rsquared)
162
163     gamma_pooled = np.mean(pooled_results) if pooled_results else

```

```

    np.nan
152     se_pooled = np.std(pooled_results) / np.sqrt(len(
pooled_results)) if pooled_results else np.nan
153     r2_pooled = np.mean(pooled_r2) if pooled_r2 else np.nan
154
155     # Regressione individuale
156     individual_results = []
157     t_values = []
158     individual_r2 = []
159     for col in price_cols:
160         company = col.split('_')[0]
161         df_period[f'{company}_Vt_rt'] = df_period[f'{company}
_LogTurnover_Triangular_Detrended_Hampel'] * df_period[f'{
company}_Return']
162
163         X_indiv = sm.add_constant(df_period[[f'{company}_Vt_rt']
+ [f'{col}_rt' for col in dummies.columns]].shift(1)).dropna()
164         y_indiv = df_period[f'{company}_Return'].dropna().loc[
X_indiv.index]
165
166         if not X_indiv.empty and not y_indiv.empty:
167             model_indiv = sm.OLS(y_indiv, X_indiv).fit(cov_type='
HC3')
168             individual_results.append(model_indiv.params[1])
169             t_values.append(model_indiv.tvalues[1])
170             individual_r2.append(model_indiv.rsquared)
171
172     gamma_mean = np.mean(individual_results) if
individual_results else np.nan
173     negative_gammas = len([g for g in individual_results if g <
0])
174     significant_gammas = len([t for t in t_values if t < -1.64])
175     t_mean = np.mean(t_values) if t_values else np.nan
176     r2_mean = np.mean(individual_r2) if individual_r2 else np.nan
177
178     results.append([
179         period_name,
180         gamma_ew, se_ew, r2_ew,
181         gamma_pooled, se_pooled, r2_pooled,
182         gamma_mean, negative_gammas, significant_gammas, t_mean,
r2_mean
183     ])
184
185     columns = [
186         'Period',
187         'EW ', 'SE EW ', 'R EW',
188         'P ', 'SE P ', 'R Pooled',
189         ' ', 'Negative s ', 'Significant s ', 't ', 'R Mean'
190     ]
191
192     results_df = pd.DataFrame(results, columns=columns)

```

```
193 print(results_df)
```

Listing 3.6: Codice Python per l'analisi dei dati

Codici Modello Teorico

In questa parte vengono esposti i codici che sono stati implementati per l'ottenimento del modello simulativo.

Si rinviene che la scelta di alcuni parametri risulta fondamentale ed i risultati possono variare in base ai parametri scelti e al numero di simulazioni svolte.

Simulazioni di Z_t , parametrizzazione di p_Z e p_0 e simulazione andamento di D_t e F_t

```
1 #!/usr/bin/env python
2 # coding: utf-8
3
4 # In[1]:
5
6
7 import numpy as np
8 import matplotlib.pyplot as plt
9
10 # Parametri iniziali
11 Z_bar = 3 # Valore iniziale di Z_t
12 lambda_val = 0.25
13 sigma_F = 0.01
14
15 # Funzione per simulare il processo stocastico di Z_t e b_t con
    il nuovo Z_t
16 def simulate_process(alpha_Z, T=3000, a=3, omega=0.0005, R
    =1.00004):
17     # Simulazione del processo stocastico tilde_Z_t
18     tilde_Z_t = np.zeros(T)
19     tilde_Z_t[0] = 0 # Partiamo da 0 per tilde_Z_t
20
21     # Processo per Z_t con Z_bar come valore di partenza
22     Z_t = np.zeros(T)
23     Z_t[0] = Z_bar # Valore iniziale di Z_t
24
25     # Simulazione di tilde_Z_t e Z_t
26     for t in range(1, T):
27         sigma_Z_squared = lambda_val**2 * (R - alpha_Z)**2 /
sigma_F**2
28         sigma_Z = np.sqrt(sigma_Z_squared)
29         u_Z_t = np.random.normal(0, sigma_Z) # Termine
stocastico u_Z,t
30
31         # tilde_Z_t segue il processo stocastico alpha_Z *
tilde_Z_t-1 + u_Z_t
32         tilde_Z_t[t] = alpha_Z * tilde_Z_t[t-1] + u_Z_t
```

```

33
34     # Z_t = Z_bar + tilde_Z_t
35     Z_t[t] = Z_bar + tilde_Z_t[t]
36
37     # Calcolo di b_t utilizzando la nuova equazione
38     b_t = (Z_t * (1 - omega) * a) / (1 - (Z_t * omega))
39
40     return Z_t, b_t
41
42 # Valori di alpha_Z da simulare
43 alpha_Z_values = [0, 0.25, 0.5]
44
45 # Creazione dei grafici per ogni alpha_Z
46 for alpha_Z in alpha_Z_values:
47     # Simulazione per il valore corrente di alpha_Z
48     Z_t, b_t = simulate_process(alpha_Z)
49
50 #-----FINE SIMULAZIONE ANDAMENTO Z_T e b_t
51
52
53 # In[2]:
54
55
56 # PARAMETRIZZAZIONE DI p_Z e p_0
57
58 # Parametri iniziali
59 R = 1.01
60 r = 0.01
61 a = 3
62 omega = 0.0005
63 T = 3000 # Numero di periodi
64 Z_bar = 3 # Valore iniziale di Z_t
65 lambda_val = 0.25
66 sigma_F = 0.01
67
68 # Funzione modificata per simulare Z_t e calcolare p_Z e p_0
69 def simulate_modified_pZ_p0(alpha_Z):
70     # Simulazione del processo stocastico tilde_Z_t e Z_t
71     tilde_Z_t = np.zeros(T)
72     tilde_Z_t[0] = 0 # Partiamo da 0 per tilde_Z_t
73     Z_t = np.zeros(T)
74     Z_t[0] = Z_bar # Partiamo da Z_bar per Z_t
75
76     # Liste per salvare p_Z e p_0 nel tempo
77     p_Z_values = []
78     p_0_values = []
79
80     # Simulazione del processo stocastico per Z_t e calcolo
81     # dinamico di p_Z e p_0
82     for t in range(1, T):
83         # Calcolo della varianza sigma_Z^2 e sigma_Z_star_sq

```

```

83     sigma_Z_squared = lambda_val**2 * ((R - alpha_Z)**2 /
sigma_F**2)
84     sigma_Z_star_sq = (R - alpha_Z)**2 / (4 * sigma_F**2)
85     sigma_Z = np.sqrt(sigma_Z_squared)
86     u_Z_t = np.random.normal(0, sigma_Z) # Termine
stocastico u_Z,t
87
88     # tilde_Z_t segue il processo stocastico alpha_Z *
tilde_Z_t-1 + u_Z_t
89     tilde_Z_t[t] = alpha_Z * tilde_Z_t[t-1] + u_Z_t
90
91     # Z_t = Z_bar + tilde_Z_t
92     Z_t[t] = Z_bar + tilde_Z_t[t]
93
94     # Calcolo di p_Z e p_0 per ogni t
95     p_Z = -((R - alpha_Z) / (2 * sigma_Z_squared)) * (1 - np.
sqrt(1 - (sigma_Z_squared / sigma_Z_star_sq)))
96     p_0 = ((1 - alpha_Z) * p_Z * Z_bar) / r
97
98     # Salva i valori di p_Z e p_0 per il grafico
99     p_Z_values.append(p_Z)
100    p_0_values.append(p_0)
101
102    return np.array(p_Z_values), np.array(p_0_values)
103
104 # Simulazioni per alpha_Z = 0, 0.25, 0.5
105 alpha_Z_values = [0, 0.25, 0.50]
106
107 for alpha_Z in alpha_Z_values:
108     p_Z_values, p_0_values = simulate_modified_pZ_p0(alpha_Z)
109
110 #-----FINE- PARAMETRIZZAZIONE DI p_Z e p_0
111
112
113 # In[3]:
114
115
116 #SIMULAZIONE ANDAMENTO DEL DIVIDENDO E DEL VALORE FONDAMENTALE
CON sigma_S^2 e sigma_ ^2
117
118 # Parametri iniziali
119 R = 1.01
120 r = 0.01
121 a = 3
122 omega = 0.0005
123 T = 5000 # Numero di periodi
124 Z_bar = 3 # Valore iniziale di Z_t
125 D_bar = 0.01 # dividendo medio
126 D_0 = 0.01 # Dividendo iniziale
127 lambda_val = 0.25
128 sigma_F = 0.01

```

```

129 alpha_D = 1
130
131 # Simulazione casuale dei pesi per sigma_S^2 e sigma_u^2
132 def random_sigma_u():
133     weight_sigma_S = np.random.uniform(0, 1) # Peso casuale per
sigma_S^2
134     weight_sigma_epsilon = 1 - weight_sigma_S # Peso
complementare per sigma_u^2
135     sigma_S_squared = (R - alpha_D)**2 * sigma_F**2
136     sigma_epsilon_squared = (R - alpha_D)**2 * sigma_F**2 / R**2
# Volatilita in base a sigma_u^2
137     return (weight_sigma_S * sigma_S_squared) + (
weight_sigma_epsilon * sigma_epsilon_squared)
138
139 # Funzione per simulare i valori di F_t, D_t e P_t con il nuovo
Z_t
140 def simulate_process_modified(alpha_Z):
141     # Simulazione del processo stocastico tilde_Z_t
142     tilde_Z_t = np.zeros(T)
143     tilde_Z_t[0] = 0 # Partiamo da 0 per tilde_Z_t
144
145     # Processo per Z_t con Z_bar come valore di partenza
146     Z_t = np.zeros(T)
147     Z_t[0] = Z_bar # Valore iniziale di Z_t
148
149     # Processo per il dividendo
150     S_t = np.random.normal(0, np.sqrt((R - alpha_D)**2 * sigma_F
**2), T) # Processo stocastico per S_t
151     tilde_D_t = np.zeros(T) # Componente stocastica di D_t
152
153     # Processo per il dividendo
154     D_t = np.zeros(T)
155     D_t[0] = D_0
156
157     for t in range(1, T):
158         sigma_u_squared = random_sigma_u() # Seleziona
casualmente la varianza ponderata
159         u_D_t = np.random.normal(0, np.sqrt(sigma_u_squared)) #
Termini stocastici
160         tilde_D_t[t] = alpha_D * tilde_D_t[t-1] + u_D_t
161         D_t[t] = D_bar + tilde_D_t[t] # D_t = \bar{D} + \tilde{D}
}_t
162
163     # Calcolo della varianza sigma_Z^2 e sigma_Z_star_sq
164     sigma_Z_squared = lambda_val**2 * (R - alpha_Z)**2 /
sigma_F**2
165     sigma_Z_star_sq = (R - alpha_Z)**2 / (4 * sigma_F**2)
166     sigma_Z = np.sqrt(sigma_Z_squared)
167     u_Z_t = np.random.normal(0, sigma_Z) # Termine
stocastico u_Z,t
168

```

```

169     # tilde_Z_t segue il processo stocastico alpha_Z *
tilde_Z_t-1 + u_Z_t
170     tilde_Z_t[t] = alpha_Z * tilde_Z_t[t-1] + u_Z_t
171
172     # Z_t = Z_bar + tilde_Z_t
173     Z_t[t] = Z_bar + tilde_Z_t[t]
174
175     # Calcolo dei coefficienti p_Z e p_0
176     p_Z = -(R - alpha_Z) / (2 * sigma_Z_squared) * (1 - np.sqrt(1
- (sigma_Z_squared / sigma_Z_star_sq)))
177     p_0 = ((1 - alpha_Z) * p_Z * Z_bar) / r
178
179     # Calcolo del prezzo fondamentale F_t con p_0 e p_Z inclusi
180     F_t = (R * D_bar) / r + (R / (R - alpha_D)) * tilde_D_t + (1
/ (R - alpha_D)) * S_t
181     P_t = F_t - D_t + (p_0 + p_Z * Z_t) # Prezzo al tempo t
182
183     # Imporre il vincolo che il prezzo non sia negativo
184     P_t = np.maximum(P_t, 0)
185     F_t = np.maximum(F_t, 0)
186
187     return F_t, D_t, P_t
188
189 # Valori di alpha_Z da simulare
190 alpha_Z_values = [0, 0.25, 0.5]
191
192 for alpha_Z in alpha_Z_values:
193     F_t, D_t, P_t = simulate_process_modified(alpha_Z)
194
195 #----- FINE - SIMULAZIONE ANDAMENTO DEL DIVIDENDO E DEL VALORE
FONDAMENTALE CON sigma_S^2 e sigma_ ^2

```

Listing 3.7: Codice Python per l'analisi dei dati

Simulazione delle regressioni del modello teorico

```

1 #!/usr/bin/env python
2 # coding: utf-8
3
4 # In[ ]:
5
6
7 #SIMULAZIONE MODELLO TEORICO
8
9 import numpy as np
10 import matplotlib.pyplot as plt
11 import statsmodels.api as sm
12
13 # Parametri iniziali
14 R = 1.01
15 r = 0.01
16 a = 3

```

```

17 omega = 0.0005
18 T = 5000 # Numero di periodi
19 alpha_D = 1
20 Z_bar = 3
21 sigma_F = 0.01
22 lambda_val = 0.25
23 D_bar = 0.02
24 n_simulazioni = 100 # Numero di simulazioni
25
26 # Funzione per simulare Z_t, S_t, _D ,t e volume di trading
27 def simulate_Z_S_eps_and_volume(alpha_Z):
28     # Inizializzazione del processo Z_t e tilde_Z_t
29     Z_t = np.zeros(T)
30     tilde_Z_t = np.zeros(T)
31     Z_t[0] = Z_bar # Valore iniziale di Z_t
32
33     S_t = np.random.normal(0, np.sqrt((R - alpha_D)**2 * sigma_F
34     **2), T) # Simulazione di S_t
35     epsilon_D_t = np.random.normal(0, np.sqrt((R - alpha_D)**2 *
36     sigma_F**2 / R**2), T) # Simulazione di _D ,t
37
38     # Simulazione del processo per Z_t
39     for t in range(1, T):
40         sigma_Z_squared = lambda_val**2 * (R - alpha_Z)**2 /
41         sigma_F**2
42         sigma_Z = np.sqrt(sigma_Z_squared)
43         u_Z_t = np.random.normal(0, sigma_Z) # Termine
44         stocastico u_Z_t
45
46         # tilde_Z_t segue il processo stocastico alpha_Z *
47         tilde_Z_t-1 + u_Z_t
48         tilde_Z_t[t] = alpha_Z * tilde_Z_t[t-1] + u_Z_t
49
50         # Z_t = Z_bar + tilde_Z_t
51         Z_t[t] = Z_bar + tilde_Z_t[t]
52
53     # Calcolo dei coefficienti p_Z e p_0
54     sigma_Z_star_sq = (R - alpha_Z)**2 / (4 * sigma_F**2)
55     p_Z = -(R - alpha_Z) / (2 * sigma_Z_squared) * (1 - np.sqrt(1
56     - (sigma_Z_squared / sigma_Z_star_sq)))
57     p_0 = ((1 - alpha_Z) * p_Z * Z_bar) / r
58
59     # Calcolo di b_t
60     b_t = (Z_t * (1 - omega) * a) / (1 - (Z_t * omega))
61
62     # Calcolo di X_t^a
63     X_t_a = np.zeros(T)
64     for t in range(T):
65         X_t_a[t] = Z_t[t] / a
66
67     # Calcolo del volume di trading V_t

```

```

62     V_t = np.zeros(T)
63     for t in range(1, T):
64         V_t[t] = omega * np.abs(X_t_a[t] - X_t_a[t-1])
65
66     return Z_t, S_t, epsilon_D_t, V_t[1:]
67
68 # Funzione per calcolare i rendimenti in eccesso Q_{t+1}
69 def calculate_excess_return(alpha_Z, Z_t, S_t, epsilon_D_t, p_Z,
70 p_0):
71     Q_t_plus_1 = np.zeros(T-1)
72
73     for t in range(T-1):
74         Z_t_plus_1 = Z_t[t+1]
75         Z_t_current = Z_t[t]
76
77         # Calcolo di Q_{t+1} usando l'equazione (A.1)
78         Q_t_plus_1[t] = -r * p_0 + p_Z * (Z_t_plus_1 - R *
79 Z_t_current) +
80                             (1 / (R - alpha_D)) *
81 S_t[t+1] + (R / (R - alpha_D)) * epsilon_D_t[t+1]
82
83     return Q_t_plus_1
84
85 # Funzione per eseguire la regressione sui dati simulati per r_{t
86 +1} =      +      r_t
87 def esegui_regressione_semplice(R_t):
88     R_t_lagged = R_t[:-1] # Rendimento lagged (ritardato di un
89 periodo)
90
91     # Aggiungiamo una costante per l'intercetta (alpha)
92     X = sm.add_constant(R_t_lagged)
93
94     # Variabile dipendente: rendimento corrente
95     Y = R_t[1:]
96
97     # Eseguiamo la regressione
98     model = sm.OLS(Y, X).fit()
99
100     # Estrazione dei risultati
101     beta = model.params[1] # Coefficiente di R_{t-1} ( )
102     t_beta = model.tvalues[1] # t-statistica per
103     R_squared = model.rsquared # R^2 della regressione
104
105     return beta, t_beta, R_squared
106
107 # Funzione per eseguire la regressione sui dati simulati r_{t+1}
108 =      +      r_t +      _1 (V_t * r_t)
109 def esegui_regressione_interazione(R_t, V_t):
110     R_t_lagged = R_t[:-1] # Rendimento lagged (ritardato di un
111 periodo)
112     V_t_lagged = V_t[:-1] # Volume al tempo t-1

```

```

106     # Calcoliamo il volume detrendizzato
107     window_size = 5 # Scegliamo una finestra di 5 periodi
108     MAV_V_t = np.convolve(V_t, np.ones(window_size)/window_size,
mode='valid')
109     MAV_V_t = np.concatenate([np.zeros(window_size-1), MAV_V_t])
110     # Completiamo la serie
111     V_detrend = V_t - MAV_V_t # Volume detrendizzato
112
113     # Variabili indipendenti: il rendimento lagged e il prodotto
tra volume detrendizzato e rendimento lagged
114     X = np.column_stack([R_t_lagged, V_detrend[:-1] * R_t_lagged
])
115
116     # Aggiungiamo una costante per l'intercetta (alpha)
117     X = sm.add_constant(X)
118
119     # Variabile dipendente: rendimento corrente
120     Y = R_t[1:]
121
122     # Eseguiamo la regressione
123     model = sm.OLS(Y, X).fit()
124
125     # Estrazione dei risultati
126     beta = model.params[1] # Coefficiente di R_t-1 ( )
127     t_beta = model.tvalues[1] # t-statistica per
128     gamma_1 = model.params[2] # Coefficiente di V_t-1 * R_t-1 (
_1 )
129     t_gamma_1 = model.tvalues[2] # t-statistica per _1
130     R_squared = model.rsquared # R^2 della regressione
131
132     return beta, t_beta, gamma_1, t_gamma_1, R_squared
133
134 # Funzione per eseguire n simulazioni e calcolare i valori medi
135 def run_multiple_simulations(alpha_Z):
136     beta_list_simple = []
137     t_beta_list_simple = []
138     R_squared_list_simple = []
139
140     beta_list_interaction = []
141     t_beta_list_interaction = []
142     gamma_1_list_interaction = []
143     t_gamma_1_list_interaction = []
144     R_squared_list_interaction = []
145
146     for _ in range(n_simulazioni):
147         # Simulazione delle variabili Z_t, S_t, _D ,t e volume
di trading
148         Z_t, S_t, epsilon_D_t, V_t = simulate_Z_S_eps_and_volume(
alpha_Z)
149
150         # Calcolo di p_Z e p_0

```

```

150     sigma_Z_squared = lambda_val**2 * (R - alpha_Z)**2 /
sigma_F**2
151     sigma_Z_star_sq = (R - alpha_Z)**2 / (4 * sigma_F**2)
152     p_Z = -(R - alpha_Z) / (2 * sigma_Z_squared) * (1 - np.
sqrt(1 - (sigma_Z_squared / sigma_Z_star_sq)))
153     p_0 = ((1 - alpha_Z) * p_Z * Z_bar) / r
154
155     # Calcolo dei rendimenti in eccesso
156     R_t = calculate_excess_return(alpha_Z, Z_t, S_t,
epsilon_D_t, p_Z, p_0)
157
158     # Regressione semplice  $r_{t+1} = \beta + r_t$ 
159     beta_simple, t_beta_simple, R_squared_simple =
esegui_regressione_semplice(R_t)
160     beta_list_simple.append(beta_simple)
161     t_beta_list_simple.append(t_beta_simple)
162     R_squared_list_simple.append(R_squared_simple)
163
164     # Regressione con interazione  $r_{t+1} = \beta + \gamma_1 V_t + r_t$ 
165     beta_interaction, t_beta_interaction, gamma_1_interaction
, t_gamma_1_interaction, R_squared_interaction =
esegui_regressione_interazione(R_t, V_t)
166
167     beta_list_interaction.append(beta_interaction)
168     t_beta_list_interaction.append(t_beta_interaction)
169     gamma_1_list_interaction.append(gamma_1_interaction)
170     t_gamma_1_list_interaction.append(t_gamma_1_interaction)
171     R_squared_list_interaction.append(R_squared_interaction)
172
173     # Calcola le medie dei coefficienti per la regressione
semplice
174     beta_mean_simple = np.mean(beta_list_simple)
175     t_beta_mean_simple = np.mean(t_beta_list_simple)
176     R_squared_mean_simple = np.mean(R_squared_list_simple)
177
178     # Calcola le medie dei coefficienti per la regressione con
interazione
179     beta_mean_interaction = np.mean(beta_list_interaction)
180     t_beta_mean_interaction = np.mean(t_beta_list_interaction)
181     gamma_1_mean_interaction = np.mean(gamma_1_list_interaction)
182     t_gamma_1_mean_interaction = np.mean(
t_gamma_1_list_interaction)
183     R_squared_mean_interaction = np.mean(
R_squared_list_interaction)
184
185     # Stampa dei risultati
186     print(f"Risultati medi per alpha_Z = {alpha_Z}:")
187     print(f"--- Regressione semplice ( $r_{t+1} = \beta + r_t$ ) ---")
188     print(f"Media di beta: {beta_mean_simple}, t(beta): {
t_beta_mean_simple}")

```

```

189     print(f"Media di R-squared: {R_squared_mean_simple}\n")
190
191     print(f"--- Regressione con interazione ( $r_{t+1} = \alpha + \beta r_t + \gamma_1 V_t * r_t$ ) ---")
192     print(f"Media di beta: {beta_mean_interaction}, t(beta): {t_beta_mean_interaction}")
193     print(f"Media di gamma_1: {gamma_1_mean_interaction}, t(gamma_1): {t_gamma_1_mean_interaction}")
194     print(f"Media di R-squared: {R_squared_mean_interaction}\n")
195
196 # Valori di alpha_Z da simulare
197 alpha_Z_values = [0, 0.25, 0.5]
198
199 # Esegui la simulazione e regressione per ogni alpha_Z
200 for alpha_Z in alpha_Z_values:
201     run_multiple_simulations(alpha_Z)
202
203 #-----FINE - SIMULAZIONE MODELLO TEORICO <3

```

Listing 3.8: Codice Python per l'analisi dei dati

Bibliografia

1. Wikipedia contributors. (2023). S&P 500. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/S%26P_500
2. Beers, B. (2023). Why Do Investors Use the S&P 500 as a Benchmark? In Investopedia. Retrieved from <https://www.investopedia.com/articles/investing/082415/why-do-investors-use-sp-500-benchmark.asp>
3. Yahoo Finance contributors. (2023). What is the S&P 500 and How Does it Work? In Yahoo Finance. Retrieved from <https://finance.yahoo.com/news/what-is-the-s-p-500-123456789.html>
4. Investopedia. (2023). What Does the Dow Jones Index Measure and How Is It Calculated? Retrieved from <https://www.investopedia.com/terms/s/DJ.asp>
5. DSPyt. (2022). How to illustrate log vs simple returns. <https://dspyt.com/simple-returns-log-return-and-volatility-simple-introduction>
6. Statistics How To. (2023). Detrend Data. <https://www.statisticshowto.com/detrend-data/>
7. Arxiv.org. (2023). Effects of polynomial trends on detrending moving average analysis. Retrieved from <https://arxiv.org/abs/1505.02750>
8. Duttaviews.com. (2023). Understanding the Triangular Moving Average (TMA). <https://duttaviews.com/triangular-moving-average/>
9. Statsmodels. (2023). Stationarity and detrending (ADF/KPSS). Retrieved from <https://www.statsmodels.org/stationarity-detrending-adf-kpss>
10. Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4), 987-1007.
11. Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327.
12. Balestra, P. (1982). Dummy Variables in Regression Analysis. In *Advances in Economic Theory* (ed. M. Baranzini). Oxford: Blackwell.
13. Suits, D.B. (1957). Use of Dummy Variables in Regression Equations. *Journal of the American Statistical Association*, 52, 548-551.
14. Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). Trading Volume and Serial Correlation in Stock Returns. *The Quarterly Journal of Economics*, 108(4), 905-939.

15. Lo, A. W., & MacKinlay, A. C. (1990). An econometric analysis of nonsynchronous trading. *Journal of Econometrics*, 45(1-2), 181-211.
16. Brock, W., Lakonishok, J., & LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance*, 47(5), 1731-1764.
17. Aaron Klein. The importance of financial data collection and standardization: Highlights from the Brookings event. Brookings, 2023. Available at: <https://www.brookings.edu>
18. Corporate Finance Institute. Corporate Finance Institute, 2023. Available at: <https://corporatefinanceinstitute.com>
19. SAS Institute Inc. Analysis of Unobserved Component Models. SAS Institute Inc. (2002). Available at: <https://support.sas.com/kb/59/749.html>
20. Taylor & Francis Group. Unobserved Components Models. 2023. Available at: <https://www.taylorfrancis.com>
21. Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357-384.
22. Goldfeld, S. M., & Quandt, R. E. (1973). A Markov model for switching regressions. *Journal of Econometrics*, 1(1), 3-15.

